Endonuklease VII des Bakteriophagen T4: Sequenzierung und Überexpression

DISSERTATION

zur
Erlangung des Grades eines
Doktors der Naturwissenschaften
der
Abteilung für Biologie
an der Ruhr-Universität Bochum

vorgelegt von

JÖRG TOMASCHEWSKI

Bochum 1988

LEIDEN
BIBL.

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41543113_0021

Endonuklease VII des Bakteriophagen T4: Sequenzierung und Überexpression

DISSERTATION

zur
Erlangung des Grades eines
Doktors der Naturwissenschaften
der
Abteilung für Biologie
an der Ruhr-Universität Bochum

vorgelegt von

JÖRG TOMASCHEWSKI

Bochum 1988

Für das Zustandekommen der vorliegenden Arbeit bin ich zu besonderem Dank verpflichtet:

- Herrn Prof. Dr. W. Rüger für die Überlassung des Themas, die Bereitstellung eines Arbeitsplatzes sowie insbesondere für sein großes Interesse an dieser Arbeit und seine ständige Diskussionsbereitschaft.
- Herrn Prof. Dr. U. Winkler, an dessen Lehrstuhl diese Arbeit durchgeführt wurde.
- Frau U. Aschke-Sonnenborn für die praktische Durchführung der Reinigung von Endonuklease VII und Herrn Dipl. Biol. N. Lamm für seine Unterstützung bei der Benutzung von Computerprogrammen.
- Dem Graduiertenförderungswerk für die Finanzierung eines Forschungsaufenthaltes bei Dr. E. Kutter (Olympia, USA) und dem Deutschen Akademischen Austauschdienst (DAAD), der mir finanzielle Unterstützung für die Durchführung eines Forschungsprojektes bei Dr. D. Shub (Albany, USA) gewährte.
- Dr. D. Shub, Dr. Belfort (beide Albany, USA) und Dr. V. Tanyashin (Moskau, UdSSR) für ihr Interesse an meiner Arbeit und die vielen anregenden Diskussionen.
- Dr. S. Tabor (Boston, USA) für die Bereitstellung des pT7-Expressionssystems und die Ratschläge in der Einarbeitungsphase.
- Dr. E. Kutter für ihre große Hilfsbereitschaft während eines Forschungsaufenthalts sowie besonders für ihre Mithilfe bei dem Zustandekommen von Kooperationen mit Dr. D. Shub und Dr. L.D. Simon.
- Meinen ehemaligen "S-Blöcklern" Frau Dipl. Biol. F. Kretschmer, Frau G. Hötten, Herrn Dipl. Biol. J. Kemper, Herrn T. Koch und Herrn R. Faria für die tatkräftige Unterstützung bei den Klonierungs- und Sequenzierungsarbeiten.
- Den Mitarbeitern der AG. Molekulare Genetik für die gute Zusammenarbeit.
- Dr. B. Kemper und Herrn Dipl. Biol H. Kosak (Köln) für die Durchführung der Endonuklease VII-Aktivitätstests.
- Dr. U. Kück für die Übernahme des Korreferats.

Verzeichnis der im Text verwendeten Abkürzungen

Amp	: Ampicillin
ATP	: Adenosin-5´-triphosphat
bp	: Basenpaare
d	: Dalton
ddNTP	: 2´, 3´-Didesoxynukleosid-5´- triphosphat
dNTP	: Desoxynukleosid-5´- triphosphat
EDTA	: Ethylendiamintetraazetat (Na-Salz)
gp	: Genprodukt
GTP	: Guanosin-5´-triphosphat
HMC	: 5´-Hydroxymethylcytosin
kb	: Kilobasenpaare
kd	: Kilodalton
Km	: Kanamycin
min.	: Minute
orf	: Offener Leserahmen (open reading frame)
RF	: Replikative Form
rpm	: Umdrehungen pro Minute
RT	: Raumtemperatur
SDS	: Natriumdodecylsulfat
TEMED	: N,N,N´, N´,-Tetramethylethylendiamin
Tris	: Trishydroxymethylaminomethan
Ük	: Übernachtkultur
Xgal	: 5-Brom-4-Chlor-Indolyl-ß-D-Galaktosid

INHALTSVERZEICHNIS

		Seite
1.	EINLEITUNG	1
1.1.	Die ortsspezifische Rekombination	1
1.2.	Die allgemeine Rekombination	3
1.3.	Endonuklease VII des Bakteriophagen T4 (Genprodukt 49)	7
1.4.	Problemstellung	10
2.	MATERIAL UND METHODEN	11
2.1.	Chemikalien und Enzyme	11
2.2.	Bakterien- und Phagenstämme	11
2.3.	Nukleinsäuren	12
2.4.	Oligonukleotide	12
2.5.	Nährmedium	12
2.6.	Puffer und Lösungen	13
2.7.	DNA-Präparationen	14
2.7.1.	Isolation von DNA aus T-Phagen	14
2.7.2.	Präparation von einzelsträngiger DNA aus M13-Phagen	14
2.7.3.	Isolation von Plasmid-DNA und doppelsträngiger M13-DNA	14
2.7.3.1.	Lösungen	14
2.7.3.2.	Schnellpräparation kleiner DNA-Mengen	15
2.7.3.3.	Isolierung von größeren Mengen Plasmid-DNA	15
2.7.4.	Isolation von DNA-Fragmenten aus Agarosegelen	15
2.7.5.	Phenolextraktion	16
2.7.6.	Ethanolfällung	16
2.8.	Gelelektrophoresen	16
2.8.1.	Lösungen	16
2.8.2.	Agarosegele	18
2.8.3.	Sequenzgele	19
2.8.4.	SDS-Polyacrylamid-Gelelektrophorese (Laemmli, 1970)	19
2.9.	DNA-Sequenzierung	20
2.9.1.	DNA-Einzelstrangsequenzierung (nach Sanger et al., 1977)	20
2.9.2.	DNA-Doppelstrangsequenzierung	20

2.10. DNA-DNA-Hybridisierungen 21
2.10.1. Radioaktive Markierung einzelsträngiger M13-DNA durch Primerverlängerung 21
2.10.2. Plaque-Hybridisierung 21
2.10.3. Southern-Hybridisierung 22
2.10.4. Bestimmung der Orientierung des inserierten DNA Fragmentes in M13-Vektoren durch DNA-DNA-Hybridisierung 22
2.11. Klonierung von DNA-Fragmenten 23
2.11.1. Zerlegung von DNA mit Restriktionsenzymen 23
2.11.2. Erzeugung von DNA-Fragmenten für die "shot-gun"-Klonierung durch Ultraschall 23
2.11.2.1. Beschallen von einzelsträngiger DNA 23
2.11.2.2. Beschallen von doppelsträngiger DNA 24
2.11.3. Hydrolyse von einzelsträngiger DNA mit S1-Nuklease 24
2.11.4. Auffüllen von DNA-Fragmenten mit überstehenden 5´-Phosphatenden 25
2.11.5. Verknüpfung von DNA-Fragmenten durch T4-DNA-Ligase 25
2.11.6. Transformation von Escherichia coli-Zellen 25
2.11.6.1. Transformation von E. coli K38/pGP1-2 mit den pT7-Plasmiden 26
2.11.6.2. Transformation von E. coli K38 oder JM103 mit den pT7-Plasmiden 26
2.11.6.3. Transfektion von E. coli JM103 oder CMK603 mit M13-DNA 26
2.12. "Marker-rescue-Tests" 27
2.13. Überexpression von klonierten Genen im pT7-Expressionssystem 28
2.13.1. Überproduktion von plasmidkodierten Proteinen zu analytischen Zwecken 30
2.13.2. Maximierung der Proteinproduktion 31
2.14. Reinigung von Genprodukt 49 über eine DNA-Cellulose-Säule 32
2.15. Endonuklease VII-Aktivitätstests 32
2.16. Computerauswertung der Sequenzdaten 32

3. ERGEBNISSE 33

3.1. Sequenzierung einer 9,1 kb großen DNA-Region zwischen den Kartenpositionen 48,266 kb und 39,166 kb 33
3.1.1. Darstellung der Sequenzierungsstrategie 33
3.1.2. Bestimmung der DNA-Sequenz 34

3.2.	Auswertung der Nukleotidsequenz und Identifizierung der Gene 49, nrdC und 55	37
3.2.1.	Restriktionskarte	37
3.2.2.	Offene Leserahmen	37
3.2.3.	Identifizierung der Gene nrdC, 49 und 55	40
3.2.4.	Promotoren und Sekundärstrukturen	42
3.3.	Klonierung und Überexpression des Gens nrdC	43
3.3.1.	Beschreibung des Klonierungsablaufs	43
3.3.2.	Überexpression von Gen nrdC	45
3.4.	Klonierung und Überexpression von Gen 49	45
3.4.1.	Genetische Organisation des DNA-Bereichs von etwa 46,7 kb bis 44,7 kb auf der physikalischen Karte von T4	45
3.4.2.	Klonierung von Gen 49 (+Sekundärstruktur Sk-1) und orf 55.13 in das pT7-Expressionssystem	46
3.4.3.	Getrennte Klonierung von Gen 49 und orf 55.13 in das pT7-Expressionssystem	48
3.4.4.	Deletion der Sekundärstruktur Sk-1	53
3.4.5.	Klonierung von Gen 49 ohne Sekundärstruktur Sk-1 in das pT7-Expressionssystem	57
3.4.6.	Reinigung und teilweise Sequenzierung des von pJT7-4/41-3-15M kodierten Proteins	62
3.4.7.	Induktion von Klonen mit Gen 49 ohne Sk-1 in E. coli JM103 durch M13 mGP1-2	62
3.4.8.	Endonuklease VII-Aktivitätstests	64
3.4.9.	Strategien zur Produktionssteigerung von aktiver Endonuklease VII	66
3.4.9.1.	Konstruktion eines pT7-Vektors mit einer Struktur zur Transkriptionstermination	67
3.5.	Sequenzierung von Gen 49 sowie anliegender Regionen bei den Bakteriophagen T2 und T6	71
3.5.1.	Orf 49.1, Gen 49 und orf 55.13 der Bakteriophagen T2, T4 und T6 sind sehr homolog	73
4.	DISKUSSION	78
4.1.	Klonierung und Sequenzierung der 9,1 kb-Region des Bakteriophagen T4	78
4.2.	Offene Leserahmen auf der 9,1 kb-Sequenz	79
4.2.1.	Das Mosaikgen sunY	82
4.2.2.	Orf 49.1 ist Gen pinA	86
4.2.3.	Gen nrdC und orf 55.7	88
4.2.4.	Genprodukt 49 ist Endonuklease VII	89

4.3. Analyse der 9,1 kb-Region auf Transkriptions-Kontrollelemente 93

4.3.1. Modell einer möglichen Transkriptionskontrolle für die zwischen 48,266 kb und 39,166 kb positionierten Genstrukturen 94

4.4. Expression von Endonuklease VII 96

4.5. Vergleich der Genregionen um pinA und 49 von T2, T4 und T6 100

4.6. Vergleich der T2- und T6-Nukleotidsequenzen mit der Spleißstelle des T4-Mosaikgens sunY 101

5. ZUSAMMENFASSUNG 103

6. LITERATURVERZEICHNIS 104

7. ANHANG 115

1. EINLEITUNG

Die Molekulargenetik hat sich im Verlauf der Jahrzehnte aus der klassischen Vererbungslehre heraus zu einer eigenständigen Spezialdisziplin innerhalb der Biologie entwickelt. Sie befaßt sich mit den molekularen Grundlagen der Mutabilität, der Realisierung und Reparatur genetischer Information sowie der Rekombination und Vererbung.

Der Grundvorgang der Vererbung besteht aus der Übertragung von Erbanlagen von den Vorfahren auf die Nachkommen. Bei den Eukaryonten finden diese Prozesse in der Regel während der Mitose und Meiose statt. Demgegenüber existiert bei Prokaryonten eine größere Vielfalt von Mechanismen zur Übertragung von genetischer Information (z. B. Transduktion, Transformation und Konjugation). Eng verbunden mit der Weitergabe von Erbanlagen sind Vorgänge, die zu einer Neuverteilung genetischer Information führen. Diese Prozesse werden als Rekombination bezeichnet.

Das Enterobakterium Escherichia coli und seine Viren sind auf diesem Gebiet ein sehr gut erforschtes Modellsystem. Man unterscheidet zwischen ortsspezifischer, allgemeiner und illegitimer Rekombination. Auf den letzteren Rekombinationstypus soll im Rahmen dieser Arbeit nicht eingegangen werden.

1.1. Die ortsspezifische Rekombination

Das best-untersuchte Beispiel für die ortsspezifische Rekombination ist die Integration des Lambda-Phagengenoms in das E. coli-Chromosom (Nash, 1981; Weisberg & Landy, 1983). Die infizierende lineare DNA zirkularisiert innerhalb der Zelle und wird anschließend von der E. coli-Gyrase in die supercoil-Form überführt. Nur diese supercoil-DNA kann effizient in das Wirts-Chromosom integrieren. Zur Integration des λ-Genoms wird sowohl das phagenkodierte int-Protein benötigt (MW: ca. 40000 d, Kikuchi & Nash, 1979; Kotewicz et al., 1980), als auch ein wirtskodiertes Protein, IHF genannt (Integration Host Factor). Dieses Protein weist ein Molekulargewicht von etwa 20000 d auf (Nash & Robertson, 1981).

Univ. LEIDEN NL.

Ein bestimmter DNA-Abschnitt auf dem ringförmigen λ-Chromosom (att P oder POP´) weist eine starke Homologie zu einer Sequenz auf dem E. coli-Chromosom auf (att B oder BOB´). Hierbei steht "O" für die Region, wo das crossover stattfindet, "P" und "P´" sind die links und rechts dieser Region lokalisierten Sequenzen auf dem Phagengenom und "B" und "B´" die äquivalenten Sequenzen bei E. coli. Att B ist zwischen den E. coli-Genen gal und bio kartiert.

Während att B nur etwa 25 bp groß ist; erstreckt sich att P über ca. 240 bp. Beide att-Sequenzen besitzen eine 15 bp große Homologieregion ("Core-Region"), in der das int-Protein jeweils zwei um 7 Basen versetzte Einzelstrangbrüche einführt, um so kohäsive Enden zu erzeugen (s. Abb. 1).

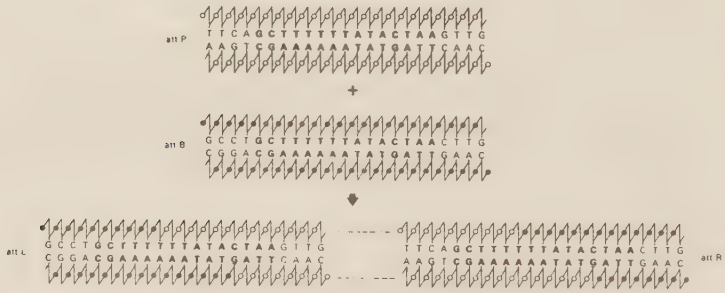


Abb. 1: Strang-Austausch bei der ortsspezifischen Rekombination des Bakteriophagen Lambda (nach Craig & Kleckner, 1986).

Die 15 bp große "Core-Region" ist in Fettdruck dargestellt.

Die Exzision der λ-DNA erfolgt über einen entgegengesetzten Reaktionsmechanismus. Diese Rekombination benötigt zusätzlich zu int und IHF das phagenkodierte Genprodukt xis. Das Protein besitzt ein Molekulargewicht von 8630 d und weist 25% Lysin- und Argininreste auf (Abremski & Gottesmann, 1982).

Xis bindet spezifisch an att L (BOP´) und att R (POB´) und katalysiert die Exzision aus dem E. coli-Chromosom (Yin et al., 1985).

Somit läßt sich für die Insertion und Exzision des λ-Chromosoms folgende Reaktionsgleichung aufstellen (nach Guerrini, 1969; Weisberg & Landy, 1983):

$$\text{POP´(att P)} + \text{BOB´(att B)} \underset{\text{gp: int, IHF, xis}}{\overset{\text{gp int, gp IHF}}{\rightleftharpoons}} \text{BOP´(att L)} + \text{POB´(att R)}$$

Es sind eine Reihe weiterer "lambdoider" Phagen bekannt, die ebenfalls über eine ortsspezifische Rekombination ins Wirts-Chromosom integrieren können und über Gene homolog zu int und xis verfügen, wie zum Beispiel P22, Ø80 und 434 (Leong et al., 1985; Leong et al., 1986).

1.2. Die allgemeine Rekombination

Bei der allgemeinen Rekombination wird die DNA zweier homologer DNA-Doppelstränge durch Bruch und Vereinigung von Einzelsträngen ausgetauscht. Dabei möglicherweise entstehende Lücken innerhalb eines Einzelstrangs werden unter Benutzung des anderen DNA-Moleküls als Matrize durch die DNA-Polymerase aufgefüllt. Zwischenstadium der Rekombination ist die Entstehung einer Heteroduplex-Region, wenn beide an der Rekombination beteiligten DNA-Doppelstränge zwar in weiten Abschnitten identisch sind, sich aber durch einzelne Basen voneinander unterscheiden. Hierdurch weist die Heteroduplex-Struktur Fehlpaarungen auf, die durch Reparaturenzyme beseitigt werden ("mismatch-repair").

Die Vorstellungen vom Verlauf der allgemeinen, generellen Rekombination sind wesentlich durch Untersuchungen an Bakterien und Viren geprägt worden. So wurden bei dem Bakteriophagen T4 zwei DNA-Doppelstrang-Moleküle nachgewiesen, die in einem kurzen Abschnitt durch Wasserstoffbrückenbindungen verknüpft waren (Anraku & Tomiszawa, 1965). Solche Strukturen können als Rekombinations-Intermediate angesehen werden.

Bei dem Bakterium E. coli sind die für die allgemeine Rekombination notwendigen Enzymaktivitäten zuerst erforscht worden. Schon 1965 wurde eine E. coli-Mutante isoliert, bei der die Rekombinationsrate nach Konjugation oder Transduktion reduziert war (Clark & Margulies, 1965). Die Mutante wurde als recA$^-$ bezeichnet. Das für die allgemeine Rekombination wichtige Genprodukt besitzt zwei unterschiedliche Enzymaktivitäten:

1) Eine Proteasefunktion, welche bei der Regulation der Expression bestimmter Gene beteiligt ist, indem Repressormoleküle gespalten werden. Diese Aktivität scheint für die Rekombination nicht notwendig zu sein, da manche recA$^-$-Mutanten, welche eine reduzierte Protease-Aktivität aufweisen, keine Reduzierung der Rekombinationsrate zeigen (Morand et al., 1977; Roberts & Roberts, 1981).

2) Das recA-Protein bewirkt die Paarung homologer DNA-Bereiche und katalysiert einen Einzelstrangtransfer in die Rezipienten-DNA-Doppelhelix. Es kommt hierbei zur Ausbildung eines "D-Loops", siehe Abb. 2. Genprodukt recA kann sowohl an Einzelstrang- als auch an Doppelstrang-DNA binden.

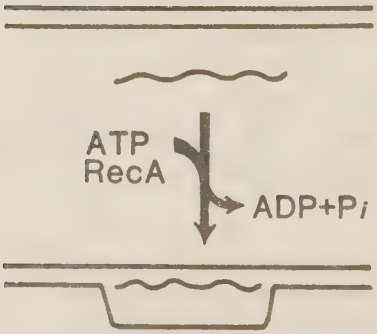


Abb. 2: Bildung eines "D-Loops" durch die Aktivität von Genprodukt recA.

Wahrscheinlich muß für diese Reaktion die einzelsträngige DNA vollständig mit recA-Protein bedeckt sein. Ist die DNA stattdessen mit Einzelstrang-Bindeproteinen wie das E. coli-SSB-Protein oder Genprodukt 32 von T4 assoziiert, reduziert sich die benötigte Menge an gp recA (Mc Entee et al., 1980; Shibita et al., 1980). Das Molekulargewicht des Proteins beträgt 37800 d (Horii et al., 1980; Sancar et al., 1980).

Bestimmte DNA-Sequenzen können die recA-abhängige Rekombination stimulieren. Diese Abschnitte werden Chi-Sequenzen genannt und weisen die Basenabfolge GCTGGTGG auf (Stahl, 1979). Chi-Sequenzen kommen auf dem E. coli-Chromosom etwa alle 5000 bp vor, werden zumindest in vitro von Endonuklease V von E. coli erkannt und in der Umgebung dieser Sequenzen geschnitten. Hierbei wird aber nur der Strang, welcher die Chi-Sequenz besitzt, hydrolytisch gespalten. Die Aktivität von Endonuklease V bewirkt also einen Einzelstrangbruch (Ponticelli et al., 1985; Taylor et al., 1985).

Endonuklease V setzt sich aus den Genprodukten recB, recC und recD zusammen und wird deshalb auch als recBCD-Nuklease bezeichnet (Telender-Muskavitz & Linn, 1982; Amundsen et al., 1986). Das recBCD-Enzym weist ebenso wie gp recA verschiedene katalytische Aktivitäten auf:

1) An doppelsträngiger DNA besitzt es eine Exonuklease-Aktivität und kann beide DNA-Stränge voneinander trennen.

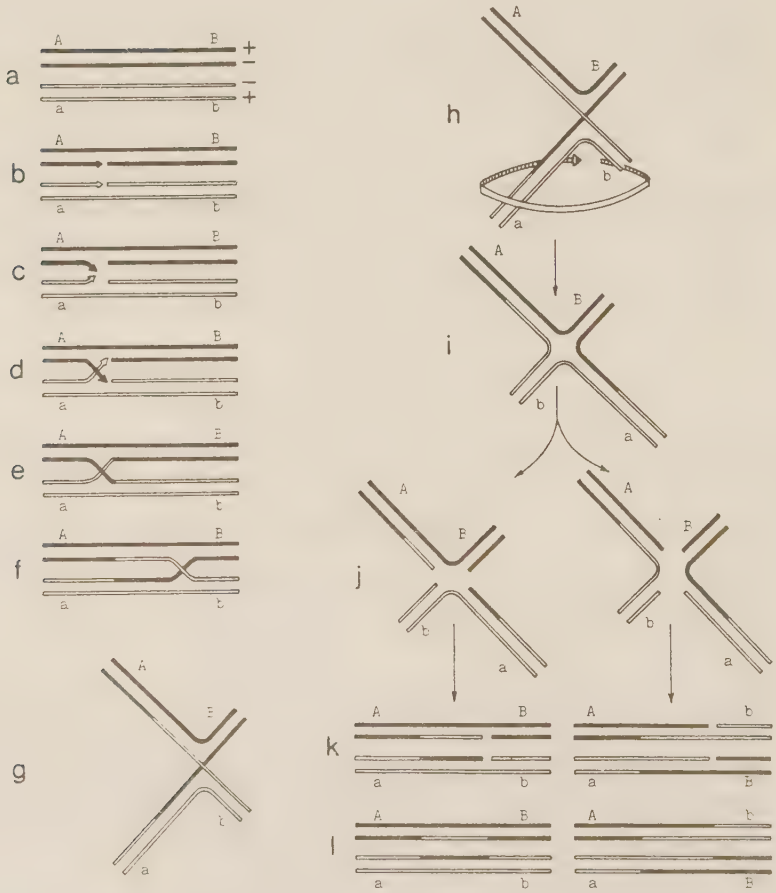
2) Einzelsträngige DNA wird durch eine endo- und exonukleolytische Aktivität gespalten (Smith et al., 1984; Taylor & Smith, 1985).

Neben Genprodukt recA und recBCD kommen in E. coli weitere Rekombinationssysteme vor. So wird bei Endonuklease V^--Mutanten noch eine restliche Rekombinationsrate gemessen. Hierfür scheint unter anderem das recF-System verantwortlich zu sein. Auch das recF-System benötigt Genprodukt recA.

Für den Ablauf der Rekombination gibt es eine Vielzahl von Vorstellungen. Vielen Modellen sind folgende drei Stadien gemeinsam: Zuerst werden in die doppelsträngigen DNA-Moleküle Einzelstrangbrüche eingeführt. Anschließend kommt es zur Ausbildung einer Heteroduplex-Region durch Strang-Wanderung ("branch-migration"). Schließlich werden die überkreuzten DNA-Stränge geschnitten und wieder zusammenligiert (s. Abb. 3).

Abb. 3: Ein Modell der allgemeinen Rekombination mit einer Holliday-Struktur als Rekombinations-Intermediat (nach Dressler & Potter, 1982).

Die beiden DNA-Stränge sind farblich unterschiedlich dargestellt. Nach der Einführung von Einzelstrangbrüchen (b) und Überkreuzung zweier Einzelstränge (c-d) mit anschließender Ligation (e) kommt es zur Strangwanderung ("branch migration", f).
Abb. g-i zeigen die Isomerisierungsreaktionen der Holliday-Struktur. Die Auflösung durch Endonukleasen (j) kann transversal oder longitudinal erfolgen und führt somit zur Ausbildung zweier unterschiedlicher Rekombinanten (k). Die Einzelstrangbrüche werden anschließend durch DNA-Ligase verbunden (l).



Den ersten beiden Schritten können E. coli-Enzyme zugeordnet werden. Einzelstrangbrüche werden durch Endonuklease V erzeugt. Auch DNA-Synthese an dem 3´-Ende eines durch Einzelstrangbruch entstandenen DNA-Moleküls und anschließender Verdrängung des ursprünglichen DNA-Strangs durch DNA-Polymerase I führt zur Erzeugung eines freien DNA-Einzelstrangs. Die Ausbildung einer Heteroduplex-DNA kann durch Genprodukt recA katalysiert werden (Cox & Lehmann, 1981). Durch sterische Umlagerung (Isomerisierung) entsteht eine Kreuzfigur (Holliday-Struktur), die durch enzymatischen Einschnitt und Neuverknüpfung gegenüberliegender DNA-Stränge aufgelöst wird (s. Abb. 3).

Für diesen letzten Schritt, das Schneiden der kreuzförmigen, "kruziformen" DNA konnte allerdings bis heute kein E. coli-Enzym nachgewiesen werden. Arbeiten am Bakteriophagen T4 dagegen haben zur Entdeckung einer Endonuklease geführt, die in vitro spezifisch Rekombinationsstrukturen schneiden kann. Dieses Enzym soll im folgenden Kapitel näher beschrieben werden.

1.3. Endonuklease VII des Bakteriophagen T4 (Genprodukt 49)

Werden E. coli-Zellen mit einer T4-Mutante infiziert, die einen Defekt in Gen 49 aufweist, so wird eine hochmolekulare DNA synthetisiert, die einen viel größeren Sedimentationskoeffizienten besitzt, als die DNA, welche nach Infektion mit einem Wildtyp-Phagen resultiert (Frankel et al., 1971). Aufgrund des hohen Molekulargewichts wird die DNA auch als "VFS-DNA" (very fast sedimenting-DNA) bezeichnet (Kemper & Janz, 1976). VFS-DNA ist hochverzweigt (Kemper & Brown, 1976). Die Verzweigungen sind oft in der Nachbarschaft teilweise gefüllter Phagenköpfe lokalisiert. Anscheinend wird der Einbau der T4-DNA in die Köpfe zwar initiiert, kann aber aufgrund der Verzweigungen nicht zu Ende geführt werden.

Diese Vermutung wurde von zwei Arbeitsgruppen unabhängig voneinander bestätigt. Frankel et al. (1971) wiesen im Rohextrakt von T4-infizierten Zellen ein Enzym nach, welches in vitro VFS-DNA endonukleolytisch spaltet. Der Test beruht darauf, daß ^{3}H-markierte VFS-DNA mit Rohextrakten aus T4-infizierten Zellen inkubiert wird. Da Endonuklease VII VFS-DNA degradiert, können nach hochtouriger Zentrifugation die radioaktiv-markierten Abbauprodukte im Überstand nachgewiesen werden.

Luftig et al. (1971) haben dagegen einen in vivo-Test durchgeführt. Werden E. coli-Zellen bei 41,5°C mit der thermosensitiven Gen 49⁻-Mutante tsC9 infiziert, so kommt es zur Bildung von VFS-DNA und teilweise gefüllten Phagenköpfen. Bei Absenkung der Temperatur auf 25°C wird die verzweigte DNA aufgelöst und in die Phagenköpfe verpackt. Das Experiment wurde so interpretiert, daß aufgrund des bei der hohen Temperatur defekten Genprodukts 49 VFS-DNA resultiert.

Woraus setzen sich die Verzweigungen zusammen? Versuche von Minagawa et al. (1983) ergaben, daß VFS-DNA aus doppelsträngiger DNA mit kurzen einzelsträngigen Abschnitten besteht. Solche Strukturen werden auch bei Rekombinationsstrukturen beobachtet (Broker, 1973). Vermutungen, daß es sich bei Endonuklease VII von T4 um ein Enzym handelt, welches spezifisch Rekombinations-Intermediate auflösen kann (Kemper & Brown, 1976) wurden durch zahlreiche in vitro Tests-bestätigt.

Teilgereinigte Endonuklease VII aus T4-infizierten Zellen führt spezifisch an Holliday-Strukturen Einzelstrangbrüche ein, welche durch DNA-Ligase geschlossen werden können (Mizuuchi et al., 1982). Es wurden eine Reihe unterschiedlicher DNA-Moleküle als Substrat für das T4-Enzym konstruiert, wie z. B. α- und χ-Strukturen. Zur Erzeugung dieser Moleküle wurden zwei pBR322-Derivate, welche sich durch zwei verschiedene Antibiotika-Resistenzgene und singuläre Schnittstellen für jeweils ein anderes Restriktionsenzym unterscheiden, in E. coli kotransformiert. Die Selektion auf Zellen, welche beide Plasmide enthalten, erfolgte auf Doppelresistenzplatten. Aufgrund der sehr ähnlichen Nukleotidabfolgen in großen Bereichen der Plasmide kommt es in den Zellen manchmal zu einer Rekombination zwischen den beiden Vektoren. Diese Rekombinationsstrukturen, welche die Form einer "8" aufweisen ("figure-eight"), können durch mehrmalige Zentrifugation in einem Cäsiumchlorid-Dichtegradienten angereichert werden (Mizuuchi et al., 1982). Werden die "figure-eight"-Moleküle mit derjenigen Restriktionsendonuklease inkubiert, welche nur in eines der beiden Plasmide einen Schnitt einführt, resultieren α-Strukturen. Nach endonukleolytischer Spaltung mit dem Enzym, welches eine Schnittstelle in dem jeweils anderen Plasmid einführt, lassen sich χ-Moleküle isolieren.

Werden diese Moleküle mit Endonuklease VII inkubiert, kommt es zur Auflösung der Rekombinationsstrukturen, was sowohl elektronenmikroskopisch, als auch aufgrund der veränderten Laufeigenschaften während der Gelelektrophorese nachgewiesen werden kann.

Endonuklease VII schneidet auch an Palindromen in supercoil-Plasmiden (Lilley & Kemper, 1984). Zum Beispiel schneidet das Enzym an einem 11 bp großen Palindrom des Plasmids pBR322. Haarnadelstrukturen können somit als ein Modellsystem für Rekombinationsstrukturen angesehen werden.

Das T4-Enzym ist nicht die einzige Nuklease, welche in vitro Holliday-strukturen schneidet. Auch Endonuklease I des Bakteriophagen T7 katalysiert diese Reaktion (de Massy et al., 1985). Interessanterweise scheint darüberhinaus die Integration des Phagen Lambda in das E. coli-Chromosom über eine Holliday-Struktur als Intermediat abzulaufen. Hsu & Landy (1984) konstruierten kruziforme DNA aus den att P- und att B-Sequenzen. Gereinigtes int-Protein kann ohne das IHF- und xis-Protein diese Strukturen auflösen.

Zwischen Endonuklease VII von T4 und Genprodukt int gibt es allerdings mehrere Unterschiede: Das int-Protein schneidet nur diejenigen Holliday-Strukturen, welche die att-Sequenzen aufweisen, während die Aktivität des T4-Enzyms sequenzunabhängig zu sein scheint. Weiterhin besitzt nur gp int eine Ligasefunktion. Schließlich bewirkt das int-Protein wahrscheinlich sowohl die Ausbildung als auch die Auflösung von Holliday-Strukturen, während Endonuklease VII nur die Auflösung von Rekombinations-Intermediaten zu katalysieren scheint.

Vor kurzem wurde auch bei Saccharomyces cerevisiae ein Enzym mit ähnlicher Aktivität wie die Endonuklease VII von T4 entdeckt (Symington et al., 1984).

Während an der Aufklärung der enzymatischen Aktivitäten von Endonuklease VII intensiv gearbeitet wird, ist über die Genorganisation und den molekularen Aufbau des Proteins nur wenig bekannt. Die Untersuchungen werden zum Teil dadurch erschwert, daß Endonuklease VII aus T4-infizierten Zellen nur in geringer Menge isoliert werden kann. Auch das gereinigte Protein ist bisher nicht auf Polyacrylamidgelen identifiziert worden.

Aus Kreuzungsanalysen ist bekannt, daß Gen 49 zwischen Gen 55 und Gen nrdC kartiert (Tessman & Greenberg, 1972). Gen 55 kodiert für ein Protein, das an der Modifikation der E. coli-RNA-Polymerase beteiligt ist, so daß das Enzym die Transkription nicht mehr an Wirtspromotoren sondern an späten T4-Promotoren initiiert (Kassavetis & Geiduschek, 1984). Dieses Gen wurde 1985 sequenziert (Gram & Rüger, 1985). T4-Thioredoxin ist ein kleines Protein, welches als Wasserstoffdonator bei der Ribose-Reduktion beteiligt ist und von Gen nrdC kodiert wird. Weitere Gene sind in diesem Abschnitt des T4-Genoms nicht kartiert, er erscheint somit als "genetisch leer".

1.4. Problemstellung

Das Genom des Bakteriophagen T4 umfaßt etwa 166000 Basenpaare (Kim & Davidson, 1974) und ist für eine Reihe von Restriktionsendonukleasen kartiert worden. Die Region zwischen den Genen 55 und nrdC ist auf der physikalischen Karte von T4 bei etwa 39 kb bis 48 kb lokalisiert.

Eine Sequenzierung der Region ist aus mehreren Gründen interessant. Zum einen sollte die Kenntnis der Nukleotidabfolge von Gen 49 die Klonierung des Gens in einen Expressionsvektor erheblich erleichtern. Nach Überexpression von Gen 49 sollen Messungen der endonukleolytischen Aktivität des synthetisierten Proteins Aufschluß darüber geben, ob Endonuklease VII alleine von Gen 49 kodiert wird, oder ob zur Erlangung der spezifischen Aktivität noch weitere von T4 kodierte Proteine benötigt werden.

Darüberhinaus sollte mit Hilfe der DNA-Sequenz eine Teilcharakterisierung des Genomabschnittes zwischen Gen 55 und nrdC erfolgen, um Aufschluß über eine mögliche Kodierungskapazität dieser "genetisch leeren" Region zu erhalten. Daß hier wahrscheinlich weitere Genstrukturen kartieren, lassen Arbeiten von Gott et al. (1986) und Simon & Fay (1982) vermuten. So wurde nachgewiesen, daß zwischen 44 kb und 49 kb eine intervenierende Sequenz positioniert ist. Diese Struktur ist somit das dritte Intron, welches in T4-Genen gefunden wurde. Die Kenntnis der Nukleotidsequenz ist Voraussetzung zur Festlegung der genauen Position des Mosaikgens. Deshalb sollte nach Sequenzierung der Region um Gen 49 in Zusammenarbeit mit den Arbeitsgruppen von Dr. Shub und Dr. Belfort die exakte Spleißstelle des Mosaikgens ermittelt werden. Weiterhin wird in der Nachbarschaft von Gen 49 ein für einen Proteinase-Inhibitor kodierendes Gen ("pin" = proteolysis inhibition) vermutet. Auch hier sollte die Kenntnis der genauen Basenabfolge nachfolgende Arbeiten zur Charakterisierung des Genprodukts wesentlich erleichtern.

Neben der Überexpression von Gen 49 stellte sich die Aufgabe, Gen nrdC in einen Expressionsvektor zu klonieren und das Genprodukt in großen Mengen anzureichern. Gen 55 wurde bereits von H. Gram überexprimiert und sollte deshalb in dieser Arbeit nicht kloniert werden.

Schließlich sollten Hybridisierungs-Experimente zeigen, ob auch die zu T4 nahe verwandten Phagen T2 und T6 eine zu Gen 49 homologe Genstruktur besitzen und die Gene gegebenenfalls sequenziert und verglichen werden.

2. MATERIAL UND METHODEN

2.1. Chemikalien und Enzyme

Wenn nicht anders angegeben, wurden die Chemikalien der Firmen Merck (Darmstadt), Serva (Heidelberg) und Baker (Groß-Gerau) verwendet.

Restriktionsendonukleasen, T4-DNA-Ligase und DNA-Polymerase (Klenow-Fragment) wurden von BRL-Gibco (Neu-Isenburg) bezogen. Alkalische Phosphatase, Lysozym und S1-Nuklease stammen von Boehringer Mannheim.

2.2. Bakterien-und Phagenstämme

Stamm	Genotyp	Bezugsquelle
E. coli BA	Wildtyp	Stammsammlung
E. coli JM103	Δ(lac, pro), supE44, thi, strA, endA, sbcB, F´traD36, proAB$^+$, lacI^q ZΔM15.	"
E. coli CMK 603	thi1, thr1, leuB6, lacY1, tonA21, supE44, r_k^-, m_k^+, F´	"
E. coli K38/pGP1-2	HfrC(λ)	S. Tabor, Boston
E. coli HMS174/pT7-3	recA1, hsdR$^-$	"
E. coli HMS174/pT7-4	recA1, hsdR$^-$	"
E. coli HMS174/pT7-5	recA1, hsdR$^-$	"
E. coli HMS174/pT7-6	recA1, hsdR$^-$	"
E. coli K38	HfrC(λ)	"
E. coli Cr63	su_I^+	Stammsammlung
E. coli K803	gal, met, supE, r_k^-, m_k^-	"
T4	Wildtyp	"
T4 49$^-$ (amE727)	Gen 49$^-$	J. Wiberg, Rochester
M13 mp18		U. Kück, Bochum
M13 mp19		"
M13 mGP1-2	T7-RNA-Polerymerase-Gen	S. Tabor, Boston

2.3. Nukleinsäuren

Lachs-DNA	Serva, Heidelberg
DNA isoliert aus T4 alc 536	Labor
DNA isoliert aus T2	Labor
DNA isoliert aus T6	Labor
Poly(dA)·poly(dT)	Boehringer, Mannheim

2.4. Oligonukleotide

Sequenzprimer M13	5´d(GTAAAACGACGGCCAGT)3´	M. Gait, Cambridge
Markierungsprimer	5´d(CGTGCTTTCCTCGTT)3´	D. Brown, Cambridge
Deletionsprimer Gen 49	5´d(ATCTATTTAAGGTAAGG)3´	H. Gram, Basel
pT7-Sequenzprimer	5´d(CGACTCACTATAGGGAGACC)3´	H. Gram, Basel
pT7-Reversprimer	5´d(CTCATGTTTGACAGCTTATC)3´	H. Gram, Basel
MB 47	5´d(CAGTCCAGTCAGTTCCCG)3´	M. Belfort, Albany

Anm.: Der pT7-Reversprimer hybridisiert nicht an das Plasmid pT7-4.

2.5. Nährmedium

TBY-Medium	10g Bacto Trypton (Difco)
	5g Hefe Extrakt (Difco)
	5g NaCl
	ad 1000 ml H_2O
PAGFe-Medium	50mM Na_2HPO_4
	22mM KH_2PO_4
	19mM NH_4Cl
	28mM Glukose
	2mM $MgSO_4$
	0,01% $FeCl_3$

Supplement für *E. coli* K38: 20µg/ml Thiamin

2.6. Puffer und Lösungen

Denhardt-Lösung 50x	Ficoll 1% (w/v) Polyvinylpyrrolidin 1% (w/v) Rinderserumalbumin 1% (w/v)
PEG-Lösung	20g PEG 6000 14.6g NaCl ad 100 ml H_2O
S1-Puffer 5x	1 M NaCl 250mM K-Azetat pH4,6 5mM $ZnSO_4$ 2,5 % Glycerin
SET 20x	3 M NaCl 0,4 M Tris HCL pH8,0 0,02 M Na_2-EDTA
SSC 20x	3 M NaCl 0,3 M Trinatriumcitrat
TE-Puffer	10mM Tris HCL pH7,9 0.1mM Na_2-EDTA
TM-Puffer 10x	100mM Tris HCl pH 8.0 50mM $MgCl_2$
WF	10mM Tris HCl pH8,0 1mM $MgCl_2$ 17mM NaCl

Für die hydrolytische Spaltung von DNA durch Restriktionsendonukleasen wurden die vom Hersteller angegebenen Puffer verwendet.

2.7. DNA-Präparationen

2.7.1. Isolation von DNA aus T-Phagen

Die Präparation von T2-, T4-, und T6-DNA erfolgte nach der von Hahn (1986) für T4 beschriebenen Methode.

2.7.2. Präparation von einzelsträngiger DNA aus M13-Phagen

1,5 ml TBY-Medium wurden mit 15µl einer Übernachtkultur (ÜK) von *E. coli* JM103 und einem M13-Einzelplaque beimpft und für 6-7 h bei 37°C unter guter Belüftung bebrütet. Die Bakterien wurden durch 10-minütige Zentrifugation bei 10000 x g sedimentiert, der Überstand mit 175 µl PEG-Lösung gemischt und 15 min. bei Raumtemperatur inkubiert. Die Phagen wurden durch eine weitere 10-minütige Zentrifugation gefällt und in 100 µl TE-Puffer resuspendiert.

Die Isolation der einzelsträngigen DNA erfolgte durch eine Phenolextraktion. Die DNA wurde anschließend gefällt und in 30µl TE-Puffer aufgenommen. Aus 1,5 ml Kulturmedium konnten zwischen 10 - 15 µg einzelsträngige M13-DNA isoliert werden.

2.7.3. Isolation von Plasmid-DNA und doppelsträngiger M13-DNA

Die hier beschriebenen Präparationen wurden in Anlehnung an die von Birnboim und Doly (1979) publizierte Methode durchgeführt.

2.7.3.1. Lösungen

Lösung I	50 mM Glucose
	25 mM Tris-HCl pH 8,0
	10 mM Na_2-EDTA
	5 mg/ml Lysozym
Lösung II	0,2 N NaOH
	1 % SDS
Lösung III	5 M Kaliumazetat-Lösung, pH 4,8

2.7.3.2. Schnellpräparation kleiner DNA-Mengen

Diese Methode wurde zur schnellen Isolierung von rekombinanter, doppelsträngiger DNA (Klonanalyse) verwendet.

1,5ml ÜK wurden durch 3-minütige Zentrifugation bei 10000 x g sedimentiert. Das Zellpellet wurde in 100µl Lösung I aufgenommen und für 15 min. bei Raumtemperatur inkubiert. Nach der Zugabe von 200µl Lösung II und einer 5-minütigen Inkubation auf Eis wurde der Ansatz mit 150µl Lösung III vermischt und für 1 Stunde auf Eis belassen. Nach einer 15-minütigen Zentrifugation wurde der Überstand in ein zweites Eppendorf-Reaktionsgefäß überführt und die DNA gefällt. Das Sediment wurde in 50µl TE-Puffer aufgenommen und 20µl zur Klonanalyse auf ein 0,9% Agarosegel aufgetragen.

2.7.3.3. Isolierung von größeren Mengen Plasmid-DNA

1l von sich in der stationären Phase befindenden E.coli-Zellen wurden durch 10-minütige Zentrifugation bei 10000 x g sedimentiert, das Pellet in 8ml Lösung I resuspendiert und für 30 Minuten bei Raumtemperatur inkubiert. Nach Zugabe von 16ml Lösung II erfolgte eine 10-minütige Inkubation bei $0^{o}C$. Anschließend wurde der Ansatz mit 12ml Lösung III vermischt und für 1 Stunde auf Eis belassen.

Zellwandtrümmer und chromosomale DNA wurden durch eine 30-minütige Zentrifugation in einem SS34 Rotor bei 18000 rpm sedimentiert, die Plasmid-DNA aus dem Überstand mit 0,6 Volumen 2-Propanol gefällt und das Pellet in 8ml TE-Puffer aufgenommen.

Nach Zugabe von 8,5g CsCl und 0,8ml Ethidiumbromid-Lösung (5mg/ml) erfolgte eine 40-stündige Zentrifugation in einem Festwinkelrotor (60Ti) bei 40000 rpm.

Die Plasmid-Bande wurde abgetropft, die DNA-Lösung zur Entfernung des Ethidiumbromids mit wassergesättigtem n-Butanol extrahiert und zweimal gegen TE-Puffer (+10mM NaCl) dialysiert.

Die DNA wurde mit Ethanol gefällt und das Sediment in einer entsprechenden Menge TE-Puffer resuspendiert.

2.7.4. Isolation von DNA-Fragmenten aus Agarosegelen

DNA-Fragmente und ccc-Formen von M13-RF-DNA und Plasmiden wurden durch Elektroelution der DNA nach Mc Donnel et al.(1977) isoliert. Anschließend wurde die

DNA zweimal mit Phenol behandelt und mit Ethanol gefällt.

2.7.5. Phenolextraktion

Die DNA-Lösung wurde mit dem gleichen Volumen neutralisierten, wassergesättigten Phenol versetzt. Die beiden Phasen wurden vermischt und durch 5 minütige Zentrifugation getrennt. Die wäßrige, obere Phase wurde abpipettiert und die DNA mit Ethanol gefällt.

2.7.6. Ethanolfällung

Um doppel- oder einzelsträngige DNA aus einer wäßrigen Lösung zu fällen, wurde eine Endkonzentration von 0,3M Natriumazetat eingestellt, 2,5-3 Volumeneinheiten Ethanol zugegeben und der Ansatz für 20 min. bei -70°C oder 2 Stunden bei -20°C gekühlt. Durch Zentrifugation für 10 min. (Eppendorf-Zentrifuge) oder 30 min. bei 12000 x g (Sorvall Zentrifuge, Typ RC 5B, Rotor SS34) wurde die DNA sedimentiert. Das Pellet wurde zweimal mit eiskaltem 80% Ethanol gewaschen, unter Vakuum getrocknet und in einem entsprechenden Puffer aufgenommen.

2.8. Gelelektrophoresen

2.8.1. Lösungen

TBE-Elektrophoresepuffer 10x	892mM Tris HCl 25mM Na_2-EDTA 890mM Borsäure ad pH 8,3
DNA-Stoppuffer 10x	75mM Na-EDTA pH 7,0 0,5 % SDS 25mM Glycerin 20µg/ml Bromphenolblau
Ethidiumbromid-Färbelösung	5µg/ml Ethidiumbromid

40% Acrylamid-Lösung (deionisiert)	380 g Acrylamid (Serva) 20 g Bis-acrylamid (Serva) ad 1000ml H_2O
6% Sequenzgel-Mix, 0,5 x TBE	150 ml 40% Acrylamid-Stammlösung 50 ml 10 x TBE 460 g Harnstoff p.a. ad 1000 ml H_2O
6% Sequenzgel-Mix, 2,5 x TBE	150 ml 40% Acrylamid-Stammlösung 250 ml 10 x TBE 460 g Harnstoff p.a. 50 g Saccharose 50mg Bromphenolblau ad 1000 ml H_2O
Formamid-Farbmarker	100 ml Formamid 100 mg Xylene-Cyanol 100 mg Bromphenolblau 5 ml 0,2M Na_2-EDTA
Lower-Tris	1,5 M Tris HCl pH 8,8 0,4 % SDS
Upper-Tris	0,5 M Tris HCl pH 6,8 0,4 % SDS
Trenngel	15,0 % (w/v) Acrylamid 0,4 % (w/v) Bis-acrylamid 0,375 M Tris HCl pH 8,8 0,1 % (w/v) SDS 0,05 % (v/v) TEMED 0,05 % (w/v) APS

Sammelgel	3,8 % (w/v) Acrylamid
	0,27 % (w/v) Bis-acrylamid
	0,125 M Tris HCl ph 8,8
	0,1 % SDS
	0,1 % (v/v) TEMED
	0,03 % (w/v) APS
Proteingel-Laufpuffer	25mM Tris-Glycin pH 8,9
	0,1 % SDS
Protein-Probenpuffer	62,5mM Tris-HCl pH 6,8
	3 % (w/v) SDS
	10 % (v/v) Glycerin
	5 % (v/v) ß-Mercaptoethanol
	20 µg/ml Bromphenolblau
Coomassie-Blue- Färbelösung	0,25 % (w/v) Coomassie Brilliant Blue
	45,5 % (v/v) Methanol
	9 % (v/v) Essigsäure
Entfärbelösung	30 % (v/v) Methanol
	7,5 % (v/v) Essigsäure

Anm.: Wenn nicht anders angegeben, ist bei prozentualen Mengenangaben immer "(w/v)" gemeint.

2.8.2. Agarosegele

Die Auftrennung von DNA-Fragmenten erfolgte auf 0,7% bis 1,4% Agarosegelen. Die Proben wurden vor dem Auftragen mit 1/10 Volumen DNA-Stoppuffer versetzt. Als Laufpuffer diente 1xTBE, die Auftrennung erfolgte bei einer Spannung von 100V. Die Dimensionen der Gele betrugen 12 x 13 x 0,5 cm, die Laufzeiten zwischen 1,5 - 4 Stunden. Nach der Elektrophorese wurden die Gele für 15 min. mit Ethidiumbromid gefärbt.

2.8.3. Sequenzgele

Zur Auftrennung der Sequenzierungsprodukte wurden, je nach Zielsetzung, verschiedene Gelsysteme verwendet. Sequenzierungsprodukte von 20-80 bp Länge wurden auf 12%-Acrylamid-Harnstoffgelen aufgetrennt, DNA-Fragmente länger als 250 bp auf 6%-Acrylamid-Harnstoffgelen. Für alle anderen Fragestellungen wurden routinemäßig Gradienten-Sequenzgele verwendet. Bei dieser Art von Gelen wird, wie bei Biggin et al. (1983) beschrieben, ein Ionengradient von 0,5 x TBE bis 2,5 x TBE aufgebaut. Aufgrund des Ionengradienten zum unteren Gelrand hin wird eine komprimierte Bandierung im unteren Gelbereich bewirkt. Hierdurch kann man im Vergleich zu Sequenzgelen ohne Ionengradienten eine erhöhte Anzahl von DNA-Fragmenten auftrennen und einwandfrei lesen. Bei Gradientengelen war der Laufpuffer 0,5 x TBE in der oberen und 1,0 x TBE in der unteren Gelkammer. Bei den anderen beiden Gelsystemen wurde in die obere und untere Pufferkammer 0,5 x TBE gegeben. Die Abmessungen der Gelkammer sind: 40 x 20 x 0,04 cm, die Laufzeit betrug 90-120 min. bei einer konstanten Leistung von 38 Watt bei Gradientengelen, 60-90 min. bei 42 Watt für 12% Gele und 3-5 Stunden bei 43 Watt für 6% Gele.

Bei Verwendung von α^{32}P-dATP wurden die Sequenzgele mit Klarsichtfolie abgedeckt; die Autoradiographie erfolgte bei -20°C. Bei Verwendung von α^{35}S-dATP zur radioaktiven Markierung wurden die Gele mit 10% Eisessig, 10% Methanol für 20 min. fixiert und nach anschließendem Trocknen unter Vakuum auf Whatman-3MM-Papier bei Raumtemperatur für 12-72 Stunden exponiert.

2.8.4. SDS-Polyacrylamid-Gelelektrophorese (Laemmli, 1970)

Die Proteine wurden auf einem diskontinuierlichen Gelsystem aufgetrennt, das heißt, Sammelgel und Trenngel unterscheiden sich sowohl in der Polyacrylamidkonzentration als auch im pH-Wert (s. 2.8.1.).

Als Polymerisationsstarter dienten APS und TEMED. Die Dimensionen der Gele waren für das Trenngel etwa 13 x 9 x 0,15 cm und für das Sammelgel 13 x 3 x 0,15 cm. Die Dauer der Elektrophorese betrug ca. 16 Stunden bei einer konstanten Stromstärke von 5mA oder 4 Stunden bei einer Stromstärke von 20mA.

Die Proben wurden mit 1/2 Volumen Protein-Probenpuffer vermischt und vor dem Auftragen 5 min. gekocht. Nach der Gelelektrophorese wurde das Gel für 2 Stunden bei 37°C mit der Coomassie-Färbelösung angefärbt und anschließend mit der Entfärbelösung entfärbt.

Die für die Autoradiographie bestimmten Gele wurden stattdessen für mindestens 30 min. bei Raumtemperatur mit 50% Trichloressigsäure fixiert und anschließend für 30 min. in Entfärbelösung gebadet. Nach dem Trocknen auf Whatman-3MM-Papier unter Vakuum erfolgte die Autoradiographie.

2.9. DNA-Sequenzierung

2.9.1. DNA-Einzelstrangsequenzierung (nach Sanger et al., 1977)

Der Sequenzprimer (s. 2.4.) wurde an die rekombinante, einzelsträngige M13-DNA für 60 min. bei 60°C in 0,2 x TM hybridisiert und die Lösung anschließend bei Raumtemperatur für 20 min. langsam abgekühlt.

Die Sequenzierungsreaktionen wurden wie bei Biggin et al. (1983) beschrieben durchgeführt.

Es wurden Didesoxynukleotide der Firma Pharmacia (Freiburg) und Desoxynukleotide der Firma Boehringer (Mannheim) verwendet. Das α-^{32}P-dATP (spez. Aktiv. 600Ci/mmol) und α-^{35}S-dATP (spez. Aktiv. 400-600 Ci/mmol) wurden von Amersham Buchler (Braunschweig) bezogen.

Nach der Sequenzierungsreaktion wurden die Proben zur Denaturierung der DNA für 2 min. gekocht und anschließend auf ein Sequenzgel aufgetragen.

2.9.2. DNA-Doppelstrangsequenzierung

Die Sequenzierung der doppelsträngigen pT7-Plasmide wurde entsprechend dem Protokoll von H.-D. Liebig (1987) durchgeführt.

Hierzu wurde aus etwa 50 ml ÜK die Plasmid-DNA wie unter 2.7.3.3. beschrieben isoliert. Die DNA wurde aber nicht durch eine CsCl-Dichtegradienten-Zentrifugation gereinigt, sondern durch präparative Gelelektrophorese isoliert.
Die DNA wurde mit dem Restriktionsenzym BglI inkubiert und auf einem 1,2% Agarosegel aufgetrennt. Das gewünschte DNA-Fragment (etwa 1000 bp + Insert-DNA) wurde mittels Elektroelution aus der Agarose isoliert. Durch diese vorhergehende Verdauung mit der Restriktionsendonuklease BglI wurde das Plasmid linearisiert und ein Teil der Plasmid-DNA entfernt. Dadurch verringert sich die Anzahl

derjenigen DNA-Regionen, an denen der Primer unspezifisch hybridisieren kann, und somit wurde die Sequenzierungsgenauigkeit erhöht.

Etwa 0,1 - 0,5 µg des isolierten DNA-Fragmentes wurden eingetrocknet. Zur Denaturierung wurde das DNA-Pellet in 0,4 N NaOH aufgenommen und für 5 min. bei Raumtemperatur belassen. Anschließend wurden ohne zusätzliche Inkubationszeiten die folgenden Lösungen in der angegebenen Reihenfolge hinzupipettiert:

0,1 - 1,0 pMol Sequenzierungsprimer (ca. 0,6 - 6 ng DNA)
3 µl 3M Natriumazetat pH 4,8
ad 20 µl H_2O
3 Volumenanteile Ethanol abs.

Die DNA wurde gefällt, zweimal mit 80% Ethanol gewaschen, getrocknet und in 8 µl H_2O + 2 µl TM-Puffer aufgenommen.

Die nachfolgenden Sequenzierungsreaktionen entsprachen denen der DNA-Einzelstrangsequenzierung (s. 2.9.1.).

2.10. DNA-DNA-Hybridisierungen

2.10.1. Radioaktive Markierung einzelsträngiger M13-DNA durch Primerverlängerung

An etwa 1 µg einzelsträngige M13-DNA wurde ca. 1 ng Markierungsprimer bei 60°C hybridisiert. Anschließend wurden dGTP, dCTP und dTTP (Endkonzentrationen jeweils 0,25mM) und 10 µCi α-^{32}P-dATP zugegeben und die Polymerisierungs-Reaktion durch Zusatz von 0.5 Einheiten DNA-Polymerase I (Klenow-Fragment) gestartet. Nach 90 min. wurde die durch den Einbau von α-^{32}P-dATP radioaktiv markierte, partiell doppelsträngige M13-DNA in der Hybridisierungsreaktion eingesetzt.

2.10.2. Plaque-Hybridisierung

Bei der Plaque-Hybridisierung wird eine große Anzahl von M13-Klonen auf zu einem bestimmten radioaktiv markierten Phagenklon komplementäre Insert-DNA getestet.

Durch diese Technik wurden z.B. aus einer Genbank gezielt diejenigen M13-Klone ermittelt, durch deren Sequenzierung noch vorhandene "Sequenzlücken" geschlossen werden konnten.

Bei dieser Methode (modifiziert nach Benton und Davies, 1977) wurden Phagenpartikel von einer Agarplatte direkt auf eine Nylon-Membran (Pall, Dreieich) übertragen. Die Filter wurden anschließend auf mit Denaturierungslösung (1,5M NaCl, 0,5M NaOH) getränktes Whatman-3MM-Papier überführt und 5 Minuten dort belassen. Nach 10-minütigem Trocknen bei 80°C und anschließendem Überführen auf 3MM-Papier, das mit Neutralisierungslösung (3 M Natriumazetat pH 5,5) getränkt war, wurden die Filter für 2 Stunden bei 80°C getrocknet.

Zur Absättigung von potentiellen, unspezifischen DNA-Bindestellen auf den Nylon-Filtern wurden sie für 4-6 Stunden in Absättigunglösung (4 x SET, 5 x Denhardt, 100 µg/ml gescherte Lachs-DNA, 0,1 % SDS) bei 60°C inkubiert.

Die Hybridisation erfolgte nach Zugabe der radioaktiv markierten Probe in neuer Absättigungslösung bei 60°C über Nacht. Vier- bis fünfmaliges Waschen der Filter für 15 min. in 6 x SSC leitete nach dem Trocknen bei 80°C die Autoradiographie ein.

2.10.3. Southern-Hybridisierung

Der Transfer von DNA-Fragmenten, welche auf Agarosegelen aufgetrennt wurden, auf Nylon-Membranen erfolgte nach der von Southern (1975) und Wahl et al. (1979) beschriebenen Methode. Die Filter wurden von Pall (Dreieich) bezogen. Die Behandlung der Gele folgte den Angaben des Herstellers.

Die Absättigung der Filter und die Hybridisierung der radioaktiven DNA erfolgte wie bei 2.10.2. beschrieben.

2.10.4. Bestimmung der Orientierung des inserierten DNA-Fragmentes in M13-Vektoren durch DNA-DNA-Hybridisierungen

Die isolierte Einzelstrang-DNA von je zwei zu testenden M13-Klonen wurde für 1 Stunde bei 60°C in 0,2 x TM inkubiert und bei Raumtemperatur für 30 min. abgekühlt.

Waren beide inserierten DNA-Fragmente in gegenläufiger Orientierung kloniert, fand in diesem Bereich Hybridisierung statt. Diese "figure eight"-Moleküle wandern bei der Elektrophorese im Agarosegel langsamer als unhybridisierte Einzelstrang-DNA.

2.11. Klonierung von DNA-Fragmenten

2.11.1. Zerlegung von DNA mit Restriktionsenzymen

Durch Zugabe von 10-fach konzentriertem Restriktionspuffer wurde eine DNA-Lösung auf die für die jeweilige Restriktionsendonuklease optimale Ionenkonzentration gebracht. Für die Hydrolyse wurde in der Regel das Zweifache der vom Hersteller angegebenen Enzymkonzentration zugegeben. Mit Ausnahme der Enzyme Sma I und Taq I erfolgten die Inkubationen der Restriktionsansätze bei 37°C für 2 bis 16 Stunden.

2.11.2. Erzeugung von DNA-Fragmenten für die "shot-gun"-Klonierung durch Ultraschall

2.11.2.1. Beschallen von einzelsträngiger DNA

Diese Methode wurde angewendet um schnell große Mengen von zufällig verteilten, sich gegenseitig überlappenden Subfragmenten für die anschließende Sequenzierung zu erzeugen. Hierbei ist es unerläßlich, das zu fragmentierende DNA-Stück in beiden Orientierungen in M13 zu klonieren.

Die Einzelstrang-DNA eines der beiden Klone wurde für 5-50 sec. mit der Mikrospitze eines "Branson-Sonifier" bei maximaler Leistung beschallt. Die Proben befanden sich in dieser Zeit auf Eis. Diese Subfragmente wurden mit der Einzelstrang-DNA des gegenläufigen Klons hybridisiert. Da im Entwicklungszyklus von M13 jeweils nur der "+"-Strang in die Phagenhülle verpackt wird, erfolgte eine Hybridisierung nur im Bereich des inserierten Fragments. Die nicht hybridisierten einzelsträngigen DNA-Regionen wurden anschließend durch eine Inkubation mit der DNA-Einzelstrang-spezifischen S1-Nuklease hydrolysiert (s. 2.11.3.).

Die DNA wurde gefällt, gelelektrophoretisch aufgetrennt und DNA-Bruchstücke einer Größe von 400 bp - 800 bp isoliert. Anschließend wurden mit dem Klenow-Fragment der DNA-Polymerase I aus E. coli bei Anwesenheit aller 4 dNTP´s die überhängenden DNA-Enden aufgefüllt und die DNA-Fragmente in M13 kloniert (s. 2.11.5.).

2.11.2.2. Beschallen von doppelsträngiger DNA

Diese Methode zur Erzeugung von zufällig verteilten, sich gegenseitig überlappenden Subfragmenten aus einem doppelsträngigen DNA-Molekül wurde nach einer von Deininger (1983) publizierten Vorschrift im Rahmen dieser Arbeit modifiziert.

Etwa 5-30 µg doppelsträngige DNA (gereinigtes Restriktionsfragment), T4-DNA oder poly (dA)·poly(dT)-DNA wurde in 400 µl TE-Puffer aufgenommen und mit der Mikrospitze eines "Branson-Sonifier", wie unter 2.11.2.1. beschrieben, beschallt.

Abweichend von der von Deininger (1983) beschriebenen Methode wurde die DNA anschließend mit 10-100 Einheiten S1-Nuklease bei 37°C für 30 min. inkubiert. Dieser Zwischenschritt wurde eingeschoben, weil durch die S1-Nuklease einzelsträngige DNA-Bereiche hydrolysiert werden und dadurch die Effizienz der nachfolgenden Ligation erhöht wird.

Die DNA wurde gefällt, auf einem 1,2% Agarosegel aufgetrennt und, wie unter 2.11.2.1. beschrieben, weiter behandelt.

2.11.3. Hydrolyse von einzelsträngiger DNA mit S1-Nuklease

5 - 20µg DNA wurden mit etwa 50 Einheiten S1-Nuklease bei 37°C für 30 min. inkubiert. Das Enzym wurde durch eine Phenolextraktion aus dem Ansatz entfernt und die DNA gefällt.

S1-Puffer 5x : 50 mM Kaliumazetat pH 4,6
200 mM NaCl
1 mM $ZnSO_4$
0,5 % (v/v) Glycerin

2.11.4. Auffüllen von DNA-Fragmenten mit überstehenden 5´- Phosphatenden

Zur Erzeugung von glatten DNA-Enden wurden 0,5 - 5 µg DNA für 30 min. bei Raumtemperatur in Anwesenheit von je 0,2mM aller 4 dNTP´s in 0,2 x TM-Puffer mit 0,5 Einheiten Klenow-Fragment inkubiert. Das Enzym wurde anschließend durch eine Phenolextraktion inaktiviert, die DNA mit Ethanol gefällt und in einem entsprechenden Puffer resuspendiert.

2.11.5. Verknüpfung von DNA-Fragmenten durch T4-DNA-Ligase

Die Verknüpfung von DNA-Fragmenten erfolgte bei 15°C für 16 Stunden mit 1-20 Einheiten T4-DNA-Ligase in Ligationspuffer. Das Ligationsvolumen betrug zwischen 30-50 µl.

Sollten DNA-Fragmente einer definierten Größe (>250 bp) kloniert werden, so betrug das Verhältnis von dephosphoryliertem Vektor zu Insert-DNA 1:3. Bei "shot-gun"-Klonierungen, bei denen teilweise auch kleinere DNA-Fragmente im Ligationsansatz vorlagen, betrug das Verhältnis von dephosphoryliertem Vektor zu Insert-DNA dagegen etwa 2:1. Der Überschuß an Vektor-DNA sollte verhindern, daß bevorzugt kleinere DNA-Fragmente untereinander zusammen ligiert und dann mit dem Vektor-Molekül verknüpft werden. Nach der Sequenzierung des "Mischklons" käme es in der anschließenden Computerauswertung zu Problemen in der einwandfreien Zuordnung dieser Basenabfolge zu der bereits erstellten DNA-Sequenz.

5x Ligationspuffer:	250mM Tris HCL pH 8,5
	50mM $MgCl_2$
	50mM DTT
	5mM ATP

2.11.6. Transformation von Escherichia coli-Zellen

E. coli-Zellen wurden nach der Methode von Cohen et al. (1972) bzw. Dagert und Ehrlich (1979) kompetent gemacht.
Je nach Art des Vektors und der Rezipientenzellen wurden die kompetenten Zellen

unterschiedlich weiter behandelt.

2.11.6.1. Transformation von E. coli K38/pGP1-2 mit den pT7-Plasmiden

20 ml TBY (+50 µg/ml Kanamycin) wurden mit 0,2 ml ÜK E. coli K38/pGP1-2 angeimpft und bei 30°C unter guter Belüftung bis zu einer OD_{590} von 0,4 bebrütet. Die Zellen wurden durch kurze, niedertourige Zentrifugation sedimentiert, in 15 ml eiskaltem 50mM $CaCl_2$ resuspendiert und für 20 min. auf Eis gehalten. Nach erneuter Zentrifugation wurden die Zellen in 1 ml eiskaltem 50mM $CaCl_2$ aufgenommen. Zu je 200 µl Zellsuspension wurden zwischen 0,1 - 05 µg DNA gegeben und der Ansatz erneut für 45-60 min. auf Eis gestellt.

Nach einem Hitzeschock bei 34°C für 5 min. wurden die Zellen in 2,5 ml TBY (ohne Antibiotika) pipettiert und für 1,5 Stunden bei 30°C durch Rollen belüftet. In dieser Zeit wurden die plasmidkodierten Antibiotikaresistenzen exprimiert.

Anschließend wurden die Zellen auf Agarplatten, die je 50 µg/ml Ampicillin und Kanamycin enthalten ausgespatelt. Durch das Vorhandensein beider Antibiotika können nur solche Transformanden Kolonien ausbilden, die beide Plasmide enthalten (pGP1-2 und pT7). Die Agarplatten wurden für etwa 24 Stunden bei 30°C bebrütet.

2.11.6.2. Transformation von E. coli K38 oder JM103 mit den pT7-Plasmiden

Da E. coli K38 oder JM103 Zellen nicht das Plasmid pGP1-2 (T7-RNA-Polymerase-Gen unter Kontrolle des thermosensitiven cI Repressors, Kanamycin-Resistenzgen) enthalten, wurden diese Zellen bei 37°C in TBY (ohne Antibiotika) bis zu einer DD_{590} von 0,4 angezüchtet. Die Kompetenzentwicklung der Zellen erfolgte wie unter 2.11.6.1. beschrieben. Der Hitzeschock dauerte 2 Minuten bei 42°C. Nach 1,5-stündigem Bebrüten in 2,5 ml TBY bei 37°C wurden die Zellen auf Ampicillin-Resistenzplatten ausgespatelt und die Platten bei 37°C über Nacht bebrütet.

2.11.6.3. Transfektion von E. coli JM103 oder CMK603 mit M13-DNA

Die Zellen wurden wie unter 2.11.6.1. beschrieben bis zu einer OD_{590} von 0,4 angezüchtet, kompetent gemacht und die DNA zu dem Transformationsansatz gegeben.

Nach dem Hitzeschritt für 2 min. bei 42°C wurden die Zellen zusammen mit 80µl einer Übernachtkultur von E. coli JM103 oder CMK603, 10µl 2,5 % IPTG und 45µl X-gal in 3 ml 0,7% Weichagar pipettiert und ausgeplattet. Nach Bebrüten der Platten bei 37°C über Nacht konnte zwischen farblosen und blauen "Plaques" unterschieden werden (M13-Phagen lysieren ihre Wirtsbakterien nicht, sondern werden aus der Zelle ausgeschieden. Die infizierten Zellen wachsen langsamer als nicht infizierte Zellen und können in einem Bakterienrasen als "Plaques" erkannt werden).

Die M13-Derivate komplementieren in den Rezipientenzellen das defekte α-Peptid der ß-Galaktosidase (α-Komplementation). Nach der Induktion mit IPTG hydrolysiert das farblose X-gal (5-Brom-4Chlor-3-Indolyl-ß-D-Galaktosid) zu dem tiefblauen 5-Brom-4-Chlor-Indigo, was zur Bildung von blauen "Plaques" führt. Inserierte DNA bewirkt, daß der Leserahmen des phagenkodierten α-Peptids unterbrochen und damit die α-Komplementation verhindert wird. M13-Phagen mit inserierter DNA bilden deshalb farblose "Plaques" aus.

2.12. "Marker-rescue-Tests"

"Marker-rescue"-Experimente beruhen auf der in vivo Rekombination von klonierter T4-DNA mit infizierenden T4-Phagenmutanten.

Hierzu wird das klonierte T4-DNA-Fragment mit Hilfe eines Plasmid- oder Phagenvektors in E. coli-Zellen amplifiziert. Anschließend werden die Zellen mit konditional letalen T4-Mutanten infiziert; liegt das Wildtypallel auf der klonierten DNA, so kommt es durch reziproke Rekombination gegen das mutierte Allel dazu, daß ein Teil der Phagennachkommen den Genotyp des Wildtypes aufweist. Der Nachweis von Wildtyp-Nachkommen erfolgt durch Vermehrung der T4-Phagen auf einem für die T4-Mutante nicht permissiven Wirt (Mattson et al. 1977), in diesem Fall E. coli BA.

5ml TBY wurden mit 40µl einer Übernachtkultur JM103 und einem Einzelplaque des zu testenden M13-Klons beimpft. Bei einer Zelldichte von etwa $5x10^8$/ml wurde die entsprechende T4-Mutante mi einer m.o.i. von 5 zu den Zellen gegeben. Nach der Inkubation für 2 Stunden bei 37°C wurden die Zellen mit Chloroform lysiert und die Phagennachkommen auf E. coli K803 und E. coli BA getitert.

2.13. Überexpression von klonierten Genen im pT7-Expressionssystem

Das hier verwendete Expressionssystem wurde von Tabor & Richardson (1985) entwickelt und setzt sich aus zwei Plasmiden zusammen (s. Abb. 4 und 5):

1. Die Plasmide pT7-3 bis pT7-6 enthalten stromabwärts von dem späten T7-Promotor Ø10 einen Multilinker und das ß-Laktamasegen. Bei den Plasmiden pT7-3 und pT7-5 bzw. pT7-4 und pT7-6 ist der Multilinker bezogen auf den Ø10-Promotor jeweils in der gleichen Orientierung angeordnet. Das ß-Laktamasegen ist bei pT7-3 und pT7-4 in Transkriptionsrichtung und bei pT7-5 und pT7-6 entgegen der Transkriptionsrichtung von Ø10 orientiert. Die mRNA-Synthese des Plasmids erfolgt in Uhrzeigerrichtung, die von Ø10 dagegen entgegen dem Uhrzeigersinn.

2. Das Plasmid pGP1-2 besitzt das Gen 1 des Bakteriophagen T7, welches für die T7-RNA-Polymerase kodiert, unter der Kontrolle des P_L-Promotors des Bakteriophagen λ. Das Gen für den thermosensitiven cI Repressor (cI857) wird von dem lac-Promotor transkribiert. Außerdem besitzt das Plasmid ein Kanamycinresistenzgen.

Die Plasmide pT7-3 bis pT7-6 und pGP1-2 sind kompatibel. Die T7-RNA-Polymerase ist nur für T7-Promotoren spezifisch und erkennt keine E. coli-Promotoren. Im Gegensatz zur E. coli-RNA-Polymerase kann man die T7-RNA-Polymerase nicht mit Rifampicin inhibieren.

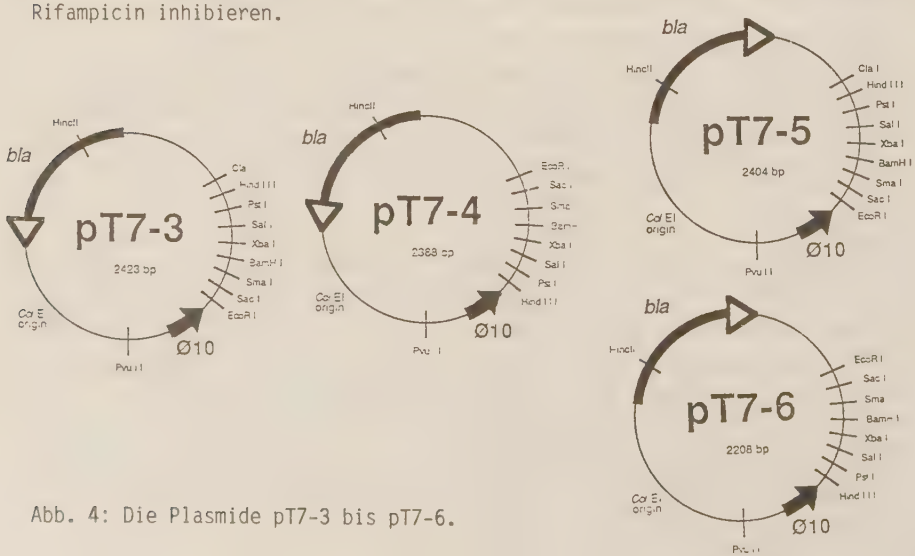


Abb. 4: Die Plasmide pT7-3 bis pT7-6.

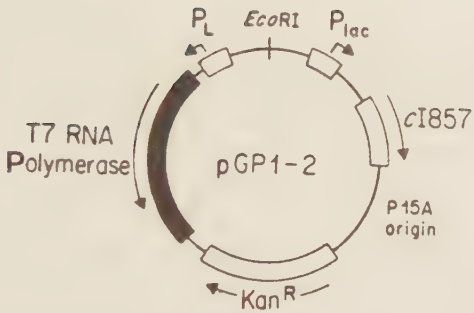


Abb. 5: Das Plasmid pGP1-2.

Soll ein bestimmtes Gen überexprimiert werden, so empfiehlt sich eine gerichtete Klonierung eines geeigneten Restriktionsfragments in Transkriptionsrichtung zum T7-Promotor. Das rekombinante Plasmid wird in den Bakterienstamm E. coli K38/pGP1-2 transformiert, welcher bereits das Plasmid pGP1-2 beherbergt. Die Selektion auf "positive" Transformanden erfolgt auf Doppelresistenzplatten (Amp./Kan.). Zellen, die sowohl das rekombinante pT7-Plasmid als auch pGP1-2 besitzen, werden bei 30°C bebrütet. Bei dieser Temperatur ist der hitzesensitive cI-Repressor aktiv und reprimiert die Transkription des Gens für die T7-RNA-Polymerase nahezu vollständig. Inaktiviert man nun den cI-Repressor durch Temperaturerhöhung auf 42°C, so wird Gen 1 transkribiert und die T7-RNA-Polymerase synthetisiert.

Eine Hemmung der E. coli-RNA-Polymerase durch Rifampicin garantiert, daß ausschließlich mRNA ausgehend vom Ø10-Promotor synthetisiert wird. Es kommt somit zur Überproduktion des von dem einklonierten Gen kodierten Proteins. Gibt man in dieser Phase radioaktiv markierte Aminosäuren zu, so wird nahezu ausschließlich das überproduzierte Protein radioaktiv markiert.

Sollen Gene überexprimiert werden, deren Genprodukte toxisch für E. coli sind, so ist es in vielen Fällen möglich, das Gen in ein pT7-Plasmid zu klonieren; aber die Transformation in E. coli K38/pGP1-2 ist für die Zellen ein letales Ereignis (eine ausführliche Diskussion dieses Sachverhalts erfolgt in Kapitel 4.4.). In solchen Fällen besteht die Möglichkeit, das rekombinante Plasmid in E. coli K38 oder JM103 zu transformieren. Die Zellen besitzen das Plasmid pGP1-2 nicht. Sind die Transformanden lebensfähig, so erfolgt die Induktion

mit dem Hybridphagen M13 mGP1-2. Dieser Phage enthält das Gen für die T7-RNA-Polymerase unter der Kontrolle des lac-Promotors kloniert. Deshalb wird zur Induktion der Synthese der phagenkodierten T7-RNA-Polymerase dem Medium IPTG zugegeben.

2.13.1. Überproduktion von plasmidkodierten Proteinen zu analytischen Zwecken

10 ml TBY (+50µg/ml Kanamycin und Ampicillin) wurden mit 0,2 ml einer Übernachtkultur von E. coli K38/pGP1-2-Zellen, die das rekombinante Plasmid enthalten, angeimpft. Bei einer OD_{590} von etwa 0,4 wurden je 5 ml Zellsuspension in einen 100 ml Erlenmeyerkolben überführt. Ein Ansatz wurde für weitere 2,5 Stunden bei 30°C bebrütet, der andere Kolben wurde zur Induktion für 25 min. bei 42°C inkubiert. Nach Zugabe von Rifampicin (Endkonzentration 200µg/ml; 20mg/ml Stammlösung in Methanol) wurden die Zellen für 2 Stunden bei 37°C bebrütet. Anschließend wurden aus beiden Ansätzen je 1ml Zellsuspension in ein Eppendorf-Reaktionsgefäß überführt, die Zellen durch eine 2-minütige Zentrifugation sedimentiert und das Zellpellet in 100µl Protein-Probenpuffer resuspendiert. Vor dem Auftragen von etwa 20-50µl auf ein Polyacrylamidgel wurden die Proben für 5 min. gekocht.

Sollten die Proteine spezifisch radioaktiv markiert werden, so wurden die Zellen in PAGFe-Medium bis zu einer OD_{590} von etwa 0,4 angezüchtet. Der Hitzeschock erfolgte bei 42°C für 25 min, wobei nach 15 min. das Rifampicin zugegeben wurde (Endkonzentration 200µg/ml). Anschließend wurde der Ansatz für 20 min. bei 30°C bebrütet. 1 ml Zellsuspension wurde in ein Eppendorf-Reaktionsgefäß überführt, 10µCi ^{35}S-Methionin zugegeben und bei 30°C für 10 min. inkubiert. Anschließend wurden die Zellen durch zweiminütige Zentrifugation sedimentiert, das Pellet in 100µl Protein-Probenpuffer resuspendiert und durch 5-minütiges Kochen aufgeschlossen. In der Regel wurden 20-30µl auf ein Polyacrylamidgel aufgetragen.

Bei der Induktion durch M13 mGP1-2 entfiel der Hitzeschock. Die Zellen wurden bei 37°C in PAGFe-Medium (+50µg/ml Ampicillin) bis zu einer OD_{590} von 0,4 angezüchtet und mit einer moi. von 10-100 mit M13 mGP1-2 überinfiziert. Gleichzeitige Zugabe von IPTG in einer Endkonzentration von 2mM induzierte den lac-Promotor und führte zur Synthese der klonierten T7-RNA-Polymerase.

Nach 25 Minuten wurde Rifampicin zugegeben und die Zellen weitere 20 Minuten bei 37°C bebrütet. Die radioaktive Markierung erfolgte genau wie für das Plasmidsystem pT7/pGP1-2 beschrieben.

Wichtig für eine effiziente radioaktive Markierung war, daß die zur Induktion verwendete M13 mGP1-2 Suspension (Titer: ca. 10^{13}/ml) keine Spuren von Methionin aufwies. Hierzu wurden die Phagen wie unter 2.7.2. beschrieben, in TBY angezüchtet, anschließend zweimal mit PEG gefällt und jedesmal in TE-Puffer aufgenommen. Außerdem wurde darauf geachtet, daß vor der Zugabe von M13 mGP1-2 Phagen die Wirtszellen nicht zu stark geschüttelt wurden, da sonst möglicherweise die F-Pili, welche für eine Infektion durch M13 benötigt werden, durch Scherkräfte beschädigt werden könnten.

2.13.2. Maximierung der Proteinproduktion

1,5 l angereichertes TBY-Medium (+ 50µg/ml Kanamycin und Ampicillin) wurden mit 30ml einer Übernachtkultur von *E. coli* K38/pGP1-2-Zellen, welche ein rekombinantes pT7-Plasmid enthalten, angeimpft. Die Kultur wurde bei 30°C bis zu einer OD_{590} von 1,0 angezüchtet. Die Induktion erfolgte durch Zugabe von 0,5 l 90°C heißem TBY-Mediums. Der Kolben wurde für 20 min. bei 42°C inkubiert und anschließend bei 37°C unter guter Belüftung für zwei weitere Stunden bebrütet. Anschließend wurden die Zellen durch 10-minütige niedertourige Zentrifugation geerntet und bei -20°C eingefroren.

Sollte ein für *Escherichia coli* schädliches Genprodukt in größeren Mengen in *E. coli* angereichert werden, so wurden 1,5 l angereichertes TBY-Medium (+ 50 µg/ml Ampicillin) mit 30 ml einer Übernachtkultur *E. coli* K38 oder JM103 Zellen, welche das rekombinante pT7-Plasmid enthalten, angeimpft und bei 37°C bei maximal 150 upm bis zu einer OD_{590} von 1,0 bebrütet. Durch eine Überinfektion mit M13 mGP1-2 (m.o.i etwa 10-100) und Zugabe von IPTG (Endkonzentration 2mM) wurden die Zellen induziert. Nach 2,5-stündiger Bebrütung bei 37°C wurden die Zellen wie oben beschrieben geerntet.

Agereichertes TBY-Medium:	50mM KPO_4 pH 7,2
	0,2 % Glycerin
	0,5 % NaCl
	1 % Hefeextrakt
	2 % Trypton

2.14. Reinigung von Genprodukt 49 über eine DNA-Cellulose-Säule

DNA-bindende Proteine können häufig durch eine Affinitätschromatographie an DNA-Cellulose gereinigt werden.

Der Aufschluß von etwa zwei Gramm induzierter Zellen erfolgte durch Zerreiben in einem Mörser nach Zugabe von 10 ml eiskaltem Aufschlußpuffer und soviel Aluminiumoxid, daß eine "pastenähnliche" Konsistenz entstand. Nach etwa 20-minütigem Aufreiben und anschließendem 30-minütigen Zentrifugieren bei 15000 upm in einem SS34-Rotor wurde der Überstand bei -20°C aufbewahrt.

Die Zusammensetzung des Aufschlußpuffers, die Herstellung der DNA-Cellulose-Säule sowie die Chromatographiebedingungen sind bei Thomas (1974) beschrieben.

2.15. Endonuklease VII-Aktivitätstests

Der Aktivitätstest wurde, wie bei Kemper und Garabett (1981) beschrieben, in Zusammenarbeit mit B. Kemper und H. Kosak durchgeführt.

Der Test beruht auf der Degradierung von Tritium-markierter VFS-DNA nach Zugabe von teilweise gereinigter Endonuklease VII.

2.16. Computerauswertung der Sequenzdaten

Alle Auswertungen der DNA-Sequenz und der hiervon abgeleiteten Aminosäure-sequenzen mit Ausnahme der Proteinhomologien erfolgten mit dem Programm von Staden (1980) auf dem Großrechner Cyber 855 des Rechenzentrums der Ruhr-Universität Bochum.

Das Proteinhomologie-Programm von Lipman und Pearson (1985) war auf einem IBM-kompatiblen Mikrocomputer implementiert.

3. ERGEBNISSE

3.1. Sequenzierung einer 9,1 kb großen DNA-Region zwischen den Kartenpositionen 48,266 kb und 39,166 kb

3.1.1. Darstellung der Sequenzierungsstrategie

Als Ausgangsmaterial zur Sequenzierung des 9,1 kb großen DNA-Abschnitts wurden folgende, leicht zu isolierende Restriktionsfragmente ausgewählt (Größenangaben nach Kutter & Rüger, 1983):

a) XbaI-14 : 2,9 kb
b) XbaI-19 : 1,3 kb
c) XbaI-18 : 1,4 kb
d) XbaI-10 : 5,3 kb

Die Abbildung 6 zeigt die sequenzierte Region und die Lage der vier genannten Restriktionsfragmente auf der physikalischen Karte von T4.

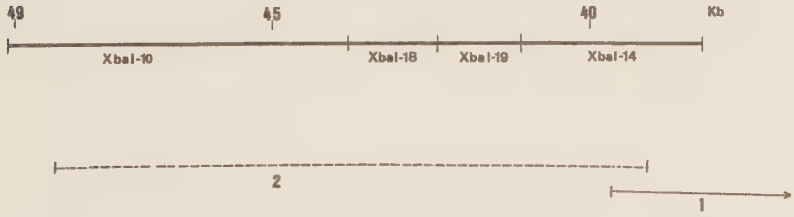


Abb. 6: Die Lage der Restriktionsfragmente XbaI-14, XbaI-19, XbaI-18 und XbaI-10 auf der physikalischen Karte des Bakteriophagen T4.

Gezeigt sind außerdem die Positionen der von Gram & Rüger (1985) publizierten Basensequenz (1) und der in dieser Arbeit erstellten Nukleotidabfolge (2).

Um diese Fragmente innerhalb eines vertretbaren Zeitraums zu sequenzieren, wurde die Strategie des "random sequencing" ausgewählt. Hierbei wird die zu sequenzierende DNA durch Ultraschall in sich gegenseitig überlappende, statistisch über das Gesamtfragment verteilte DNA-Stücke zerlegt. Die Subfragmente werden anschließend in M13 mp18 kloniert und nach der Didesoxy-Methode sequenziert. Die Einzelsequenzen werden gelesen und an ihren Überlappungsstellen mit Hilfe von geeigneten Computerprogrammen zusammengesetzt.

Im Laufe eines Sequenzierungsprojekts kann so ein Großteil der Nukleotidabfolge eines DNA-Fragments bestimmt werden. Durch den Einsatz der Plaque-Hybridisierungstechnik (s. 2.10.2.) werden anschließend diejenigen Klone aus einer Genbank ermittelt, durch deren Sequenzierung noch vorhandene Lücken in der Basenabfolge geschlossen werden können.

Durch eine gezielte Klonierung von geeigneten Restriktionsfragmenten soll anschließend über die drei XbaI-Erkennungssequenzen (XbaI-14/XbaI-19, XbaI-19/XbaI-18 und XbaI-18/XbaI-10)sequenziert werden, um so die ermittelten Nukleotidabfolgen der vier Restriktionsfragmente an ihren Schnittstellen fehlerfrei zusammensetzen zu können.

Da bereits ein Teil der DNA Sequenz des XbaI-14-Fragments (ca. 1,3kb) von Gram&Rüger publiziert wurde, sollen sich die Sequenzierungsarbeiten vor allem auf den noch nicht analysierten DNA-Bereich erstrecken. Deshalb werden bevorzugt diejenigen Klone sequenziert, die bei einer Hybridisierungsreaktion mit einer rekombinanten M13-DNA, welche ein DNA-Fragment aus der bereits publizierten Region inseriert enthält, "negativ" reagieren.

3.1.2. Bestimmung der DNA-Sequenz

Um zur Sequenzierung geeignete Subfragmente zu erzeugen, wurden zwei unterschiedliche Beschallungsmethoden angewendet:

1) Die Restriktionsfragmente XbaI-18 und XbaI-19 wurden jeweils in beiden Orientierungen in M13 mp18 kloniert. Die Fragmentierung erfolgte anschließend wie unter 2.11.2.1. beschrieben, durch Beschallung der

einzelsträngigen DNA.

2) Die Restriktionsfragmente XbaI-10 und XbaI-14 wurden direkt ohne vorhergehende Klonierung als DNA-Doppelstrang beschallt. Die Moleküle wurden auf einem Agarosegel aufgetrennt und Subfragmente einer Größe von 300 bp bis 800 bp aus dem Gel eluiert und in M13 kloniert.

Um jeweils die erhaltenen Sequenzen zweier XbaI-Restriktionsfragmente einwandfrei miteinander zu verbinden, wurden gezielt geeignete RsaI, HindIII und Eco RV Fragmente kloniert und sequenziert.

Die Auftrennung der Sequenzierungsprodukte erfolgte in der Regel auf Gradientengelen. Sollten mehr als 220 - 250 bp eines Inserts ermittelt werden, so wurden die Sequenzierungsprodukte auf 6% Polyacrylamidgelen ohne Gradient aufgetrennt. Zur fehlerfreien Bestimmung der ersten 10 - 20 bp eines Inserts wurden 12 % Polyacrylamidgele ohne Gradient eingesetzt.

Insgesamt wurde zur Erstellung der 9,1 kb großen Nukleotidsequenz die DNA von 301 Subklonen analysiert (66214 bp). Somit wurde jedes Nukleotid der Sequenz im Durchschnitt etwa siebenmal bestimmt. Auf einem Sequenzgel wurden durchschnittlich 220 bp gelesen.

Bis auf wenige Ausnahmen wurden jeweils beide DNA-Stränge sequenziert und jeder Bereich mindestens dreimal analysiert (das gilt nicht für den Teil der XbaI-14 Sequenz, der bereits von Gram & Rüger (1985) publiziert wurde).

Die Verteilung der Sequenzfolgen innerhalb der 9,1 kb Region ist der Abbildung 7 zu entnehmen.

Abb. 7: Schematische Darstellung des sequenzierten Bereichs zwischen 48,266 kb und 39,166 kb.

Gezeigt sind die Lage der offenen Leserahmen und die Restriktionsschnittstellen für Enzyme, deren Schnittpositionen bei T4 kartiert wurden (Kutter et al., 1987). Die Abkürzungen bedeuten: KpnI (K),* PstI (P), HindIII (H), 2 HindIII-Schnittstellen, 15 bp entfernt (H), HaeII (Ha), MluI (M), Eco RV (E_5), Eco RI (E), BglI (B_1), SphI (S), NdeI (N) und BglII (B_2).
Der untere Teil der Abbildung zeigt die Verteilung der 301 Einzelsequenzen, wobei ein Punkt den Startpunkt der Sequenzierung angibt.

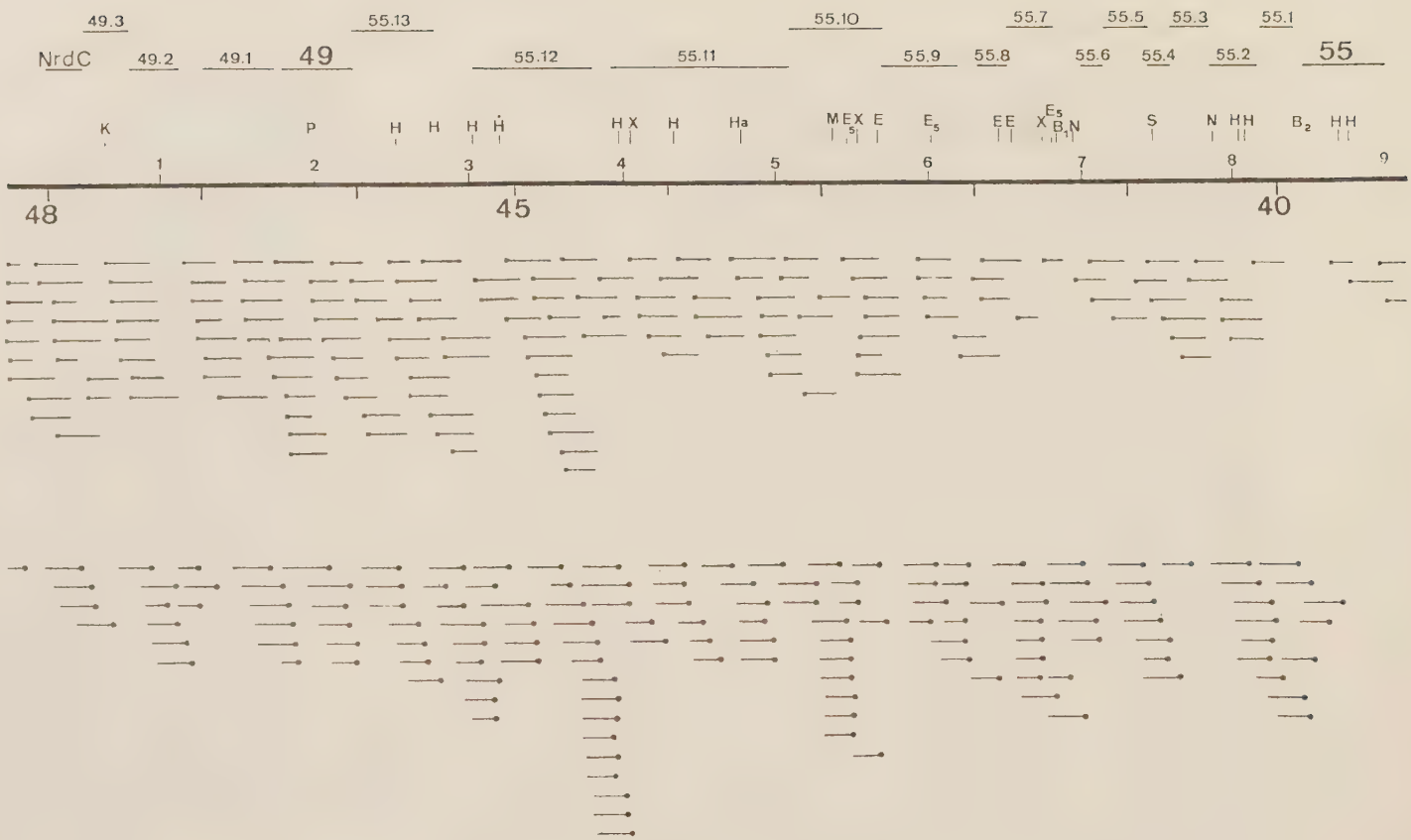

49.3
55.13
55.10
55.7
55.5
55.3
55.1
NrdC
49.2
49.1
49
55.12
55.11
55.9
55.8
55.6
55.4
55.2
55
K
P
H
H
H
Ḣ
H X
H
Ha
M E X E
E5
E E
X E5 B1 N
S
N H H
B2
H H
1
2
3
4
5
6
7
8
9
48
45
40

Abbildung 7.

Frühe T4-Promotoren lassen sich, wenn überhaupt, in M13 nur in einer Orientierung klonieren. Deshalb konnte in dieser Arbeit der DNA-Bereich um die drei frühen T4-Promotoren durch Einzelstrang-Sequenzierung nur in einer Richtung bestimmt werden. Die Sequenz des Gegenstrangs wurde mir freundlicherweise von Liebig vor einer Veröffentlichung zur Verfügung gestellt (Liebig, 1987).

Die Abbildung 34 (s. Anhang) zeigt die Nukleotidsequenz der 9,1 kb Region.

3.2. Auswertung der Nukleotidsequenz und Identifizierung der Gene 49, nrdC und 55

Die DNA-Sequenz wurde unter Einbeziehung von EDV-Programmen (Staden, 1980) auf Restriktionsschnittstellen, offene Leserahmen, Promotoren und Sekundärstrukturen untersucht.

Anschließend wurden die aus der Nukleotidsequenz abgeleiteten Genprodukte charakterisiert.

3.2.1. Restriktionskarte

In Tabelle 11 (s. Anhang) ist die mit Hilfe des Programms "cutsit" erstellte Restriktionskarte dargestellt.

3.2.2. Offene Leserahmen

Offene Leserahmen zeigen die Lage von möglichen Genstrukturen an und sind charakterisiert durch ein ATG-oder GTG-Startkodon mit einer Ribosomenbindestelle, welche 4-12 Nukleotide stromaufwärts des Starttripletts liegt (Shine & Dalgarno, 1974, Stormo et al., 1982). Begrenzt werden offene Leserahmen durch eines der drei Terminationstripletts TAA, TGA und TAG.

Nur bei einer Transkription von 48,266 kb in Richtung 39,166 kb können offene Leserahmen mit einer Größe von mehr als 120 Basen ermittelt werden.

Diese Orientierung entspricht der für "frühe" T4-Genom-Regionen beschriebenen Transkriptionsrichtung und ist damit in Übereinstimmung mit der genetischen Karte (Rüger & Kutter, 1984).

Insgesamt sind auf der 9,1 kb-Region 19 offene Leserahmen sehr eng benachbart angeordnet. Einzelne Leserahmen überlappen um ein Nukleotidpaar: sie sind durch die Sequenz ATGA miteinander verbunden, wobei das Terminationstriplett TGA mit dem Startkodon ATG des nachfolgenden Gens überlappt. Andere Leserahmen überlappen z.B um 5 Basenpaare (orf 55.6/55.5) oder 8 Basenpaare (orf 55.9/55.8).

Die Genstrukturen weisen Kodierungskapazitäten für Proteine zwischen 5000 d und 45000 d Größe auf (s. Tabelle 1). Die Abstände zwischen zwei Genen sind in der Regel nicht größer als 100 Basenpaare; in diesen Bereichen sind meist potentielle Regulationsstrukturen für die Transkription angeordnet. Eine Besonderheit bilden die intergenischen Regionen bei 2801 - 3027 und 3805 - 3911. Hierauf wird in Abschnitt 3.5.1. näher eingegangen.

Drei der offenen Leserahmen wurden als die Gene nrdC, 49 und 55 identifiziert (s. 3.2.3.).

Tab. 1: Gene und offene Leserahmen zwischen den Koordinaten 48,266 kb und 39,166 kb.

Die potentiellen Ribosomenbindestellen und Translationsstartstellen sind unterstrichen. Die Molekulargewichte der kodierten Proteine, die Anteile der positiven und negativen Aminosäuren, die Anzahl der Aminosäuren sowie der Polaritätsindex (Capaldi & Vanderkooi, 1972) sind ebenfalls aufgelistet

Für einige T4-Gene (z.B. Gen 69) wurde beschrieben, daß ein Leserahmen durch Initiation an anderen, internen Startkodons für mehr als ein Protein kodieren kann (Mosig & Macdonald, 1986). Innerhalb der 19 Genstrukturen auf der 9,1 kb-Sequenz wurden neun mögliche "Leserahmen im Leserahmen" ermittelt, welche in Tabelle 2 charakterisiert werden. Hierbei wurden nur solche Leserahmen berücksichtigt, die mit einem ATG-oder GTG-Triplett anfangen, eine Ribosomenbindestelle im Abstand von 3-12 Basenpaaren besitzen und länger als 200 bp sind.

Name	Position	Leitsequenz	% geladene As. +	% geladene As. -	Pol.Index	MW	Anzahl As
Gen nrdC	249- 509	AAGGAATAATATG	16.1	13.8	45	10038	87
orf49.3	509- 814	GGGAATACTTTAAATGA	6.9	21.6	40	11925	102
orf49.2	814-1131	AGTTGGTGGTGTGAAAAATGA	12.3	19.8	52	12565	106
orf49.1	1258-1740	TGGAGATGGGACATG	10.6	18.0	50	18795	161
Gen49	1785-2255	AAGGGTTTATTATG	19.1	13.4	54	18123	157
orf55.13	2255-2797	AAAGAGTTTAAAATGA	14.4	12.2	48	20121	181
orf55.12	3028-3801	GAGGTGAAACGATG	24.4	11.2	56	30332	258
orf55.11	3912-5102	TGGACTGGAACGAATG	14.6	11.8	46	44742	397
orf55.10	5102-5731	AGGGTTAAGCACCAATGA	22.4	8.6	55	23948	210
orf55.9	5724-6191	TTGGAAATATAAAGATG	13.5	11.5	48	18226	156
orf55.8	6312-6521	AGGTAAATGTAAAATG	11.4	8.6	39	7903	70
orf55.7	6527-6832	AAGGAATAGTTGTG	16.7	11.8	48	11706	102
orf55.6	6994-7173	GAGGAATTGATATG	20.0	8.3	40	6954	60
orf55.5	7169-7459	AAGGAAATTAAACTCCATG	28.9	17.5	62	11795	97
orf55.4	7470-7598	GAGGAAAATATG	14.0	11.6	51	5140	42
orf55.3	7602-7838	GGGAAATTTAAATG	15.2	11.4	47	9142	79
orf55.2	7855-8178	GAGGATAATGTATG	22.2	9.3	57	12712	108
orf55.1	8184-8444	AAGGTATAATTATG	5.8	12.6	36	9834	87
Gen55	8483-9037	GAGGTGAAAATG	15.7	16.0	55	21502	185

Name	Position	Leitsequenz	SD--ATG	Pol.-Index	Molekulargewicht
49,3*	626-814	AAGGT ATAGACTTTGTT GTG	12bp	43	7377d
49*	1941-2255	AGGA AAG GTG	3bp	54	12167 d
49**	1986-2255	AAGG TCAA ATG	4bp	58	10626 d
55,13*	2365-2797	GGGA CCTT ATG	4bp	44	15932 d
55,13**	2447-2797	GGA ATTATTCAG GTG	9bp	45	12906 d
55,10*	5306-5732	GGTG GAAAG ATG	5bp	57	16219 d
55,9*	5868-6191	GGA CAGTTATTCACT ATG	12bp	52	12730 d
55*	8624-9037	GGA TTAGCCATT ATG	9bp	54	16094 d
55**	8693-9037	GGA AACAAGAA ATG	8bp	55	13394 d

Tab. 2: Potentielle Genstrukturen innerhalb kodierender Sequenzen.

3.2.3. Identifizierung der Gene nrdC, 49 und 55

Die Aminosäuresequenz des T4-Thioredoxins ist bekannt (Sjöberg & Holmgren, 1972). Ein Vergleich der bereits bekannten Aminosäureabfolge mit den von den offenen Leserahmen abgeleiteten Sequenzen identifizierte den offenen Leserahmen bei Position 249 - 509 eindeutig als Gen nrdC. Es bestehen keine Unterschiede zwischen der Aminosäuresequenz und der von der Nukleotidabfolge abgeleiteten Sequenz.

Zur Identifizierung von Gen 49 wurden Kreuzungsexperimente mit der Gen 49$^-$-Mutante (amE727) durchgeführt (s. 2.12.). Das Ergebnis zeigt Tabelle 3. Durch diese Versuche wurde der offene Leserahmen bei Position 1785 - 2255 zweifelsfrei als Gen 49 bestimmt.

Gen 55 wurde bereits von Gram & Rüger (1985) sequenziert und durch "marker rescue"-Experimente identifiziert. Ein Vergleich der publizierten Sequenz mit der Nukleotidsequenz bei Position 8483 - 9037 ergab volle Übereinstimmung beider Basenabfolgen.

Klon	49^- (amE 727)	Gen 49
M13 mp18	$6,7 \times 10^{-5}$	-
M13-JT33/522	$4,7 \times 10^{-5}$	-
M13-JT33/273	$6,1 \times 10^{-4}$	+
M13-JT37/25	$4,5 \times 10^{-3}$	+
M13-JT /604	$1,0 \times 10^{-3}$	+

Tab. 3: Die Ergebnisse der Kreuzungsexperimente.

Die Zahlen geben den Anteil von Wildtyp-Rekombinanten unter der gesamten Phagennachkommenschaft an. Die Position der subklonierten DNA-Fragmente, die in diesen Experimenten verwendet wurden, ist in Abbildung 9 zu ersehen.

3.2.4. Promotoren und Sekundärstrukturen

Für eine sinnvolle Regulation der Genexpression verwendet der Bakteriophage T4 drei Klassen von Initiationsstellen für die Transkription: "frühe", "mittlere" und "späte" Promotoren. Diese drei Promotorklassen unterscheiden sich in ihren Konsensussequenzen. Auf der Sequenz der 9,1 kb-Region wurden alle drei Promotortypen ermittelt (s. Tabelle 4).

T4-"early" Promotoren (P_e1 bis P_e3)

Position	-35		-10
Konsensus	T T G A C A	- 15-21 bp-	T A T A A T
1189	T T T A C A	17 bp	T A T A A T
6935	T T T A C A	17 bp	T A T T A T
7380	T T T A C A	17 bp	T A T A A T

T4-"middle" Promotoren (P_m1 bis P_m 5)

Position	-35		-10
Konsensus	A A T G C T T	- 12-17 bp -	T A T A A T
142	T A T G C T A	14 bp	T A T T A T
2165	G A T G C T T	15 bp	T A T A A T
6065	A A T G C T T	14 bp	T A T T A T
8416	T A T G C T T	17 bp	T A T T A T
8429	A G G T C T T	12 bp	T A T A A T

T4-"late" Promotoren (P_l1 und P_l2)

Position	
Konsensus	T A T A A A T A C T A T T
1748	A A T A A A T A C C C T T
3002	T A T A A A T A T G A A T

Tab. 4: Mögliche Promotorstrukturen auf der 9,1 kb-Sequenz.

Bei der Untersuchung der aus der DNA-Sequenz abgeleiteten RNA-Sequenz auf mögliche stabile Sekundärstrukturen wurden sechs Palindrome mit einem ΔG Wert von mehr als -10 kcal/mol gefunden. Diese Sekundärstrukturen sind in Tabelle 5 dargestellt; die Bedeutung einiger dieser Sequenzen für die Regulation der Genexpression wird in Kapitel 4.3.1. diskutiert.

Position	Nukleotidsequenz	Name	ΔG/kcal
1173-1191	GAGCCGG>AAGAA<CCGGTTT	Sk-0	-10,4
1755-1779	ACCCTTATCT→ATTTA<AGGTAAGGGT	Sk-1	-13,9
2893-2929	TCCCT>GATGAATGTA<AGGGA GTA GGGTC>AAGC<GACCC	Sk-2	-14,7
2991-3026	GTTTATCCTCGTAT>AAATATGA<ATACGAGGTGAAAC	Sk-3	-12,9
5760-5783	AACGGCCCTG>GATG<TAGGGTCGTT	Sk-4	-14,2
6840-6852	GGCCC>TCT<GGGCC	Sk-5	-11,6

Tab. 5: Mögliche Sekundärstrukturen auf der Sequenz der 9,1 kb-Region.

Die freie Energie der Sekundärstrukturen (ΔG) in der RNA wurde mit dem Computerprogramm von Zuker und Stiegler (1981) berechnet. Gezeigt wird die DNA-Sequenz. Die palindromischen Sequenzen sind unterstrichen.

3.3. Klonierung und Überexpression des Gens nrdC

3.3.1. Beschreibung des Klonierungsablaufs

Aufgrund der DNA-Sequenz war bekannt, daß das Gen nrdC vollständig auf einem 0,6 kb großen HaeIII-Fragment lokalisiert ist. Dieses DNA-Fragment wurde aus zwei Gründen zuerst in M13 mp18 kloniert:

1. DNA-Sequenzierung sollte sicherstellen, daß es sich bei dem isolierten Fragment tatsächlich um das gewünschte DNA-Molekül handelt.

2. Eine vorhergehende Klonierung in M13 über stumpfe Enden in die SmaI-Schnittstelle bietet die Möglichkeit, geeignete flankierende Erkennungssequenzen für Restriktionsenzyme aus dem Multilinker auszuwählen. Nachdem das inserierte Fragment mit zwei unterschiedlichen Enzymen aus dem Multilinker herausgeschnitten wurde, kann es gezielt in einen Expressionsvektor kloniert werden ("forced cloning").

Das 0,6 kb-Fragment wurde nach Verdauung von ca. 5 µg XbaI-10-DNA aus einem 1,2% Agarosegel isoliert und in die dephosphorylierte SmaI-Schnittstelle des Vektors M13 mp18 inseriert. Der resultierende Hybridphage M13 JT36/2 enthält das 0,6 kb-Fragment mit dem vollständigen Gen nrdC. Die doppelsträngige DNA wurde isoliert und nacheinander mit den Restriktionsenzymen PstI und Eco RI inkubiert. Das 0,6 kb-Fragment wurde isoliert und in den mit PstI und Eco RI geschnittenen Expressionsvektor pT7-6 ligiert.

Durch diese gezielte Klonierung trägt das resultierende Plasmid pJT7-6/nrdC das Gen nrdC in Transkriptionsrichtung zum T7-Promotor Ø10.

Der Ablauf der Klonierung ist in Abbildung 8 schematisch dargestellt.

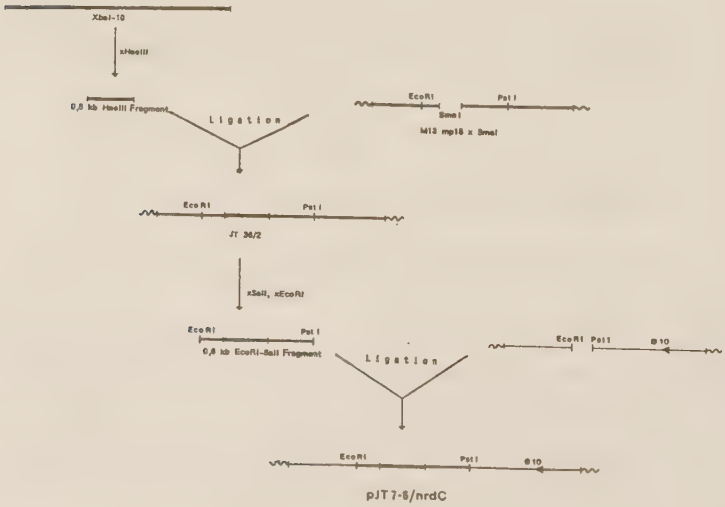


Abb. 8: Klonierung des Gens nrdC in den Expressionsvektor pT7-6.

3.3.2. Überexpression von Gen nrdC

Das rekombinante Plasmid pJT7-6/nrdC wurde in den Stamm *E. coli* K38/pGP1-2 transformiert (s. 2.11.6.1.).

Nach der spezifischen radioaktiven Markierung des T4-Thioredoxins ist nur eine Proteinbande der Größe von etwa 10 kd festzustellen (s. Abb. 13, Kap. 3.4.3.).

Die Größe des Proteins ist in guter Übereinstimmung mit dem von der Nukleotidsequenz des Gens nrdC abgeleiteten Molekulargewichts von 10038 d.

3.4. Klonierung und Überexpression von Gen 49

Zum besseren Verständnis des Klonierungsablaufs soll zuerst einführend kurz die DNA-Region um Gen 49 vorgestellt werden.

3.4.1. Genetische Organisation des DNA-Bereichs von etwa 46,7 kb bis 44,7 kb auf der physikalischen Karte von T4

In diesem DNA-Abschnitt sind die offenen Leserahmen orf 49.1, orf 55.13, orf 55.12 sowie Gen 49 lokalisiert. Weiterhin lassen sich zahlreiche potentielle Promotorstrukturen und zwei Sekundärstrukturen feststellen.

In Abbildung 9 wird die Organisation dieser Region schematisch dargestellt.

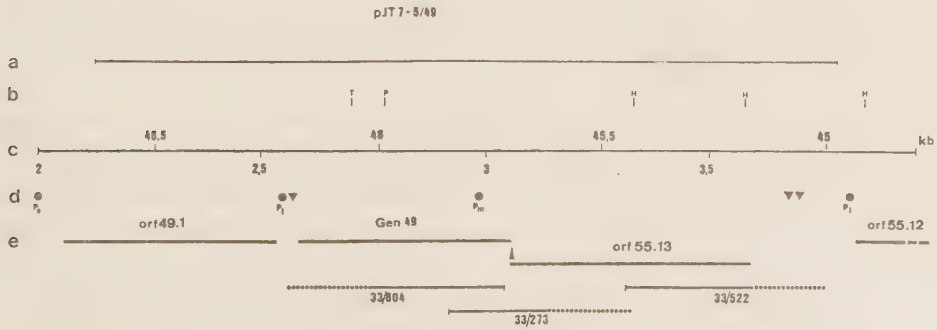


Abb. 9: Der DNA-Bereich um Gen 49.

Linie a: Position des 1,6 kb großen HpaII-PvuII-Fragments.
Linie b: Schnittstellen für Restriktionsenzyme, deren Erkennungssequenzen auf dem T4-Genom kartiert wurden (Kutter et al., 1987). Die Abkürzungen bedeuten: TaqI (T), PstI (P), HindIII (H).
Linie c: Position von Promotorstrukturen (●) und Sekundärstrukturen (▼). Frühe Promotoren sind mit P_e, mittlere Promotoren mit P_m und späte Promotoren mit P_l bezeichnet.
Linie d: Lage der offenen Leserahmen. Mit (▲) ist der ATGA-Übergang zwischen Gen 49 und orf 55.13 gekennzeichnet.
Linie e: Position der für die "marker rescue"-Versuche (s. 2.12. und Tab. 3) eingesetzten Subklone. Die durchgezogenen Linien geben den Teil des inserierten Fragments an, der sequenziert wurde. Die gepunkteten Linien zeigen die weitere Ausdehnung des Fragments. Wie bereits in Abschnitt 3.2.3. beschrieben wurde, reagierten die Klone JT33/273 und JT33/604 positiv und der Klon JT33/522 negativ. Der Klon JT37/25 trägt das 1,6 kb große, in Linie a gezeigte HpaII-PvuII-Fragments.

3.4.2. Klonierung von Gen 49 (+Sekundärstruktur Sk-1) und orf 55.13 in das pT7-Expressionssystem

Etwa 5 µg isolierte XbaI-10-DNA wurde nacheinander mit den Endonukleasen HpaII und PvuII hydrolysiert und ein etwa 1,6 kb großes DNA-Fragment isoliert (die Lage des 1,6 kb -Fragments ist in Abb. 9, Spalte a gezeigt). Auf diesem DNA-Molekül ist sowohl Gen 49 als auch der offene Leserahmen 55.13 vollständig enthalten.

Das Fragment wurde über stumpfe Enden in einen mit SmaI geschnittenen und dephosphorylierten M13 mp18-Vektor inseriert. Je nach Orientierung des Fragments in M13 wurden die Klone mit JT37/21 und JT37/25 bezeichnet. Die Kontrolle der jeweiligen Orientierungen erfolgte durch DNA-DNA-Hybridisierungen (s. 2.10.4.) und DNA-Sequenzierung.

Die Insert-DNA von JT37/21 und JT37/25 wurde mit Eco RI und SalI herausgeschnitten und in einen mit denselben Enzymen behandelten Expressionsvektor pT7-5 bzw. pT7-6 ligiert. Die resultierenden Klone wurden mit pJT7-5/49 und pJT7-6/49 bezeichnet. In Abbildung 10 ist das Klonierungsschema für pJT7-5/49 gezeigt.
Nach Induktion von pJT7-5/49 oder pJT7-6/49 ist eine drastische Überproduktion eines etwa 20 kd großen Proteins festzustellen (s. Abb. 13, Kap. 3.4.3.).

Das von der Nukleotidsequenz abgeleitete Molekulargewicht von gp 49 beträgt 18123 d, das von gp orf 55.13 dagegen 20157 d. Somit ist denkbar, daß von den Klonen pJT7-5/49 (bzw. pJT7-6/49) zwar beide Proteine überproduziert werden, aber im SDS-Gel wegen ihrer ähnlichen Größe nicht voneinander getrennt werden können und somit als eine Proteinbande erscheinen. Genausogut wäre es möglich, daß aus noch zu diskutierenden Gründen nur eines der beiden Proteine exprimiert wird.

Um diesen Sachverhalt zu klären, sollten Expressionsvektoren konstruiert werden, die entweder das Gen 49 oder orf 55.13 vollständig enthalten.

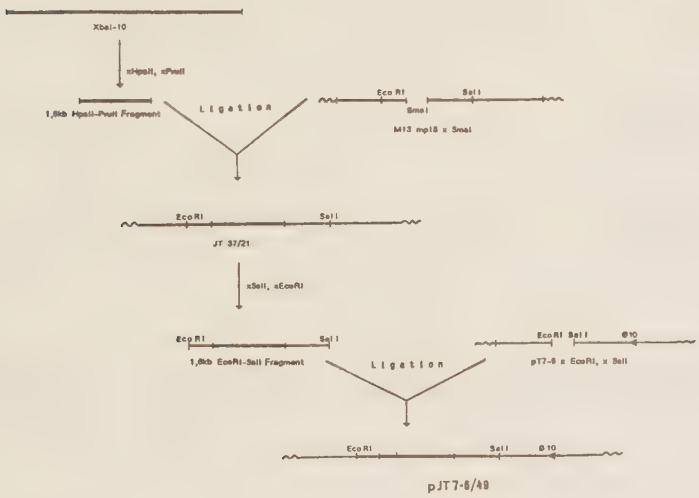


Abb. 10: Konstruktion des Klons pJT7-5/49.

3.4.3. Getrennte Klonierung von Gen 49 und orf 55.13 in das pT7-Expressionssystem

Als Ausgangmaterial dienten die beiden Klone pJT7-5/49 und pJT7-6/49. Zur Konstruktion eines Klons, welcher ausschließlich Gen 49 vollständig trägt, wurde die DNA von pJT7-5/49 mit HindIII geschnitten. Da im Insert von pJT7-5/49 zwei HindIII Schnittstellen enthalten sind, und eine weitere im Multilinker positioniert ist, wurde ein ca. 250 bp und ein ca. 210 bp großes Fragment aus dem Vektor herausgeschnitten und gelelektrophoretisch vom restlichen Plasmid abgetrennt. Nach der Selbstligation der verbliebenen Vektor-DNA wurde das Plasmid in E. coli K38/pGP1-2 transformiert.

Das Plasmid, welches neben dem vollständigen Gen 49 noch die ersten 278 bp von orf 55.13 inseriert enthält, wurde als pJT7-5/49A bezeichnet ("A"=allein). Das Klonierungsschema ist in Abbildung 11 gezeigt.

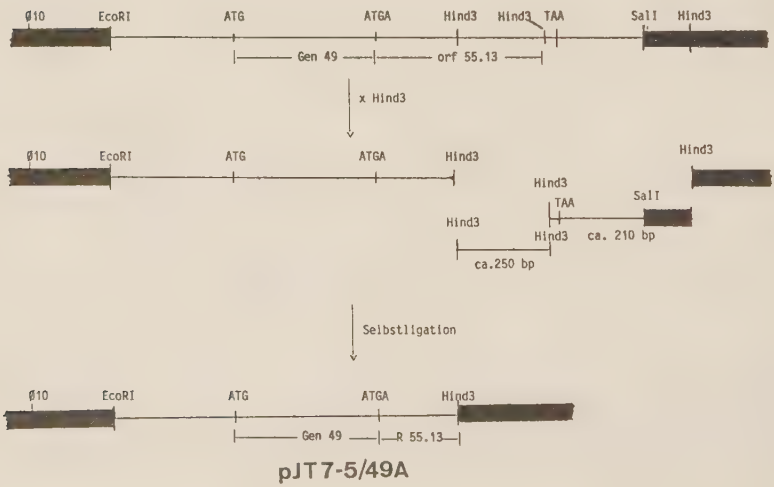


Abb. 11: Konstruktion des Vektors pJT7-5/49A aus pJT7-5/49

Zur Konstruktion eines Plasmids, welches ausschließlich orf 55.13 beinhaltet, wurde der Vektor pJT7-6/49 mit der Restriktionsendonuklease PstI verdaut. Das etwa 660 bp große DNA-Fragment wurde, wie für die Konstruktion von pJT7-5/49A beschrieben, vom restlichen Plasmid abgetrennt, dieses mit Hilfe von T4 DNA-Ligase zirkularisiert und ebenso wie das Plasmid pJT7-5/49A in E. coli K38/pGP1-2 transformiert. Das rekombinante Plasmid wurde mit pJT7-6/orf 55.13 bezeichnet. Das Klonierungsschema ist in Abbildung 12 dargestellt.

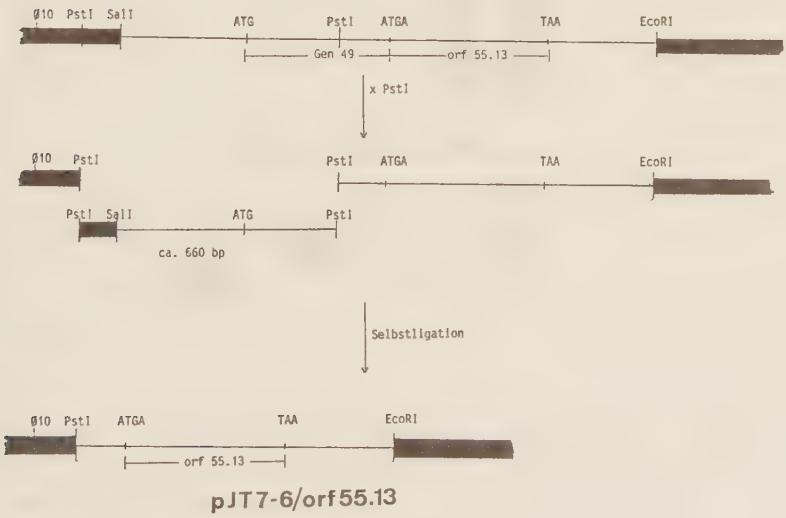


Abb. 12: Konstruktion des Vektors pJT7-6/orf 55.13 aus pJT7-6/49.

Das Plasmid pJT7-6/orf 55.13 trägt zusätzlich zum vollständigen offenen Leserahmen 55.13 noch etwa 190 bp, welche für das C-terminale Ende von Genprodukt 49 kodieren. Die für das N-terminale Ende des Proteins kodierende Genregion sowie die Ribosomenbindestelle und das ATG-Initiationskodon sind deletiert.

Die Überprüfung der Ligationsübergänge erfolgte durch DNA-Doppelstrang-Sequenzierung (s. 2.9.2.). Hierzu wurden die Plasmide pJT7-5/49A und pJT7-6/orf 55.13 aus je 40 ml Übernachtkultur isoliert und anschließend von beiden Seiten sequenziert. Die Ergebnisse zeigten, daß beide Ligationen ohne zusätzliche Basenaustausche korrekt erfolgten. Nach der radioaktiven Markierung ist deutlich die Überproduktion eines Proteins von etwa 20 kd festzustellen (s. Abb. 13).

Abb. 13: Autoradiogramm eines 15% Polyacrylamidgels.

Die Gelelektrophorese wurde nach Laemmli (1970) durchgeführt.
Die einzelnen Kanäle zeigen:
1) pJT7-5/49, uninduziert, ohne Rifampicin
2) " , induziert, ohne Rifampicin
3) " , induziert, mit Rifampicin
4) pJT7-5/49A, induziert, mit Rifampicin
5) pJT7-6/orf 55.13, induziert, mit Rifampicin
6) pJT7-6/nrdC, induziert, mit Rifampicin

Aus diesen Versuchen ist zu schließen, daß nur orf 55.13 überexprimiert wird, da nach Induktion von pJT7-5/49A ein ca. 12 kd großes Protein überproduziert wird, nicht aber ein zusätzliches 18 kd großes Protein. Bei dem 12 kd-Protein handelt es sich wahrscheinlich um das N-terminale Protein-

fragment von gp 55.13, da der Klon pJT7-5/49 die ersten 278 bp von orf 55.13 enthält. Diese Schlußfolgerung wird dadurch unterstützt, daß nach Induktion von pJT7-6/orf 55.13 genau die gleiche 20 kd große Proteinbande auftritt wie nach Induktion von pJT7-5/49, bei welchem sowohl Gen 49 als auch orf 55.13 inseriert sind.

Warum wird nur orf 55.13 überexprimiert, nicht aber Gen 49? Ein Vergleich der Ribosomenbindestellen zeigt, daß Gen 49 eine ausgeprägte Shine-Dalgarno-Sequenz besitzt, nicht aber orf 55.13 (s. Abb. 14).

Gen 49 : TATCTATTTAAGGTAAGGGTTTATT ATGT

Orf 55.13 : AGAAGCAGCTTAGAAAGAGTTTAAA ATGA

Abb. 14: Die Ribosomenbindestellen von Gen 49 und orf 55.13.

Die potentiellen Ribosomenbindestellen und das jeweilige ATG-Initiationskodon sind unterstrichen. Es ist zu beachten, daß Gen 49 mit orf 55.13 über ein ATG-Element überlappt.

Somit sollte gerade die Translation von Gen 49 effizienter sein als die von orf 55.13.

Da Gen 49 und orf 55.13 von demselben Promotor transkribiert werden (Ø10 von dem pT7-Plasmid), orf 55.13 aber stromabwärts von Gen 49 lokalisiert ist und in großer Konzentration überexprimiert wird, kann die Inhibierung von Gen 49 nicht auf der Ebene der Transkription stattfinden.

Unmittelbar stromaufwärts von Gen 49 befindet sich ein später Promotor sowie die Sekundärstruktur Sk-1. In Abbildung 15 ist diese Sekundärstruktur dargestellt. Die Anordnung der Sekundär- und Promotorstrukturen impliziert folgendes Modell für die Regulation der Genexpression von Gen 49 auf der Ebene der Translation: Während der frühen Phase des Infektionszyklus von T4 wird die Transkription vom frühen Promotor P_{e-1} aus initiiert. Diese frühe messenger RNA kann die in Abbildung 15 gezeigte Sekundärstruktur ausbilden. Da die Ribosomenbindestelle von Gen 49 im Stamm lokalisiert ist, werden im frühen Abschnitt der T4-Infektion orf 49.1, Gen 49 und orf 55.13 transkribiert,

aber nur orf 49.1 und orf 55.13 können effizient translatiert werden. Im späten Abschnitt des T4-Lebenszyklus erkennt die von T4-kodierten Proteinen modifizierte Wirts-RNA-Polymerase nicht mehr die frühen Promotorstrukturen; somit wird alle neusynthetisierte mRNA in diesem Genomabschnitt an den späten Promotoren P_{1-1} und P_{1-2} initiiert (s. Abb. 9).

Da im Falle von Gen 49 diese späte mRNA 4-5 Basen stromabwärts von P_{1-1} synthetisiert wird, kann sie nicht mehr die Sekundärstruktur Sk-1 ausbilden. Die Ribosomenbindestelle ist daher exponiert und die Synthese von Genprodukt 49 beginnt.

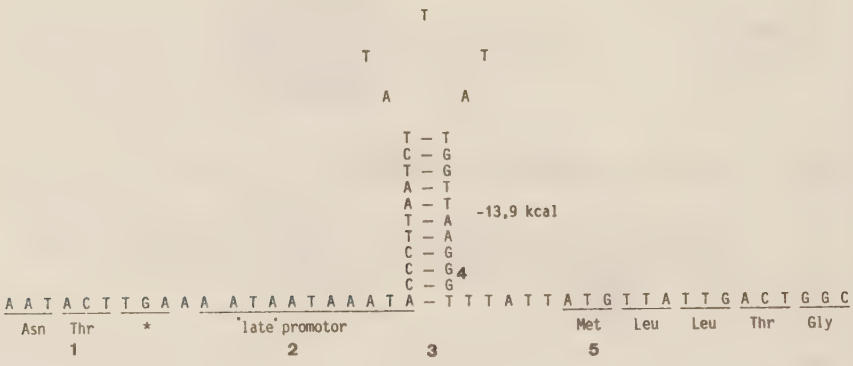


Abb. 15: Die Nukleotidsequenz der Sekundärstruktur Sk-1.

Zusätzlich zu Sk-1 (3) sind das Ende von orf 49.1 (1), der mögliche späte Promotor P_{1-1} (2), die potentielle Ribosomenbindestelle (4) und das ATG-Initiationskodon von Gen 49 (5) eingezeichnet.

Diese Art der Genregulation auf Translationsebene wurde zuerst bei den späten T4-Genen soc (small outer capsid protein) und e (Lysozym) beschrieben (Für die Expression von orf 55.12 kann ein ähnliches Modell für die Regulation der Genexpression aufgestellt werden, siehe 4.2.1.).

Das Regulationsmodell erklärt, warum im pT7-Expressionssystem nur gp 55.13, nicht aber gp 49 überproduziert wird, denn hier startet die Transkription ausgehend vom Ø10-Promotor etwa 400 bp stromaufwärts von Gen 49. Somit kann diese mRNA die beschriebene Sekundärstruktur ausbilden und die Translation von Gen 49, nicht aber die Translation von orf 55.13 inhibieren.

Folgt man diesem Modell, so ist es zur Überproduktion von Genprodukt 49 unerläßlich, die Sekundärstruktur vor Gen 49 so zu deletieren, daß die Ribosomenbindestelle nicht mehr im Stamm des Palindroms kaschiert werden kann.

3.4.4. Deletion der Sekundärstruktur Sk-1

Da die Ribosomenbindestelle von Gen 49 im Stamm der Sekundärstruktur Sk-1 lokalisiert ist, bestand das Ziel des Deletionsprojekts darin, nur die Nukleotide stromaufwärts der Shine-Dalgarno-Sequenz zu entfernen. Geeignete Restriktionsschnittstellen standen hierfür nicht zur Verfügung.

Bei einer limitierten Verdauung mit der Exonuklease Bal 31 oder DNase I besteht die Schwierigkeit, die Reaktionsbedingungen so zu wählen, daß die Hydrolyse auf etwa 10 Basen genau begrenzt ist. Es sind darüberhinaus umfangreiche Kontrollen durch DNA-Sequenzierung notwendig, um aus einer Vielzahl von Klonen diejenigen zu ermitteln, welche einerseits die Sekundärstruktur nicht mehr ausbilden können, andererseits aber noch die Ribosomenbindestelle und das ATG-Initiationskodon von Gen 49 besitzen.

Deshalb wurde im Rahmen dieser Arbeit eine Technik entwickelt, die es ermöglichen sollte, exakt die Sekundärstruktur zu deletieren. In Abbildung 16 ist das Prinzip der Methode schematisch dargestellt. Die Vorgehensweise entspricht der "site specific mutagenesis"-Technik. Im Gegensatz dazu besitzt der hier verwendete Primer aber keinen Nukleotidaustausch.

Ein Oligonukleotid mit der in Abbildung 16 gezeigten Sequenz (dick unterstrichen) wurde an die komplementäre einzelsträngige DNA des in M13 klonierten HpaII-PvuII-Fragments (Klon JT37/25) hybridisiert. Aus Gründen der Übersichtlichkeit ist in der Abbildung der zur mRNA homologe Strang gezeigt, die Hybridisierung erfolgte aber am Gegenstrang.

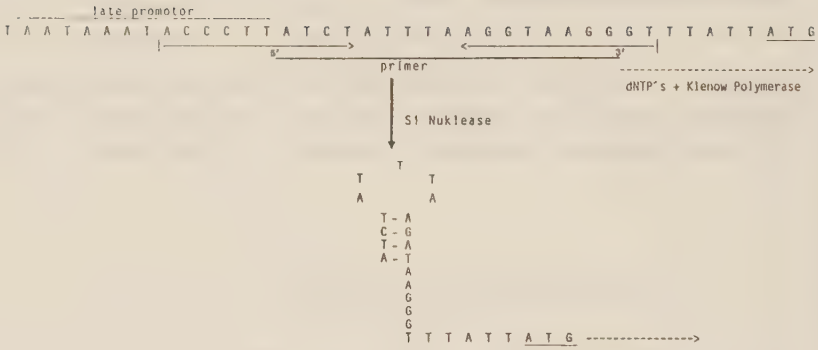


Abb. 16: Deletion der Sekundärstruktur Sk-1.

Erklärungen im Text.

Es zeigte sich, daß die Konzentration des Oligonukleotids sowie die Hybridisierungstemperatur ganz wichtige Parameter für die spezifische Hybridisierung waren. Für eine einwandfreie Anhybridisierung wurde eine Primer-Konzentration von etwa 5 nM für ca. 20 µg JT37/25 DNA eingesetzt.

Ein typischer Hybridisierungsansatz enthielt:

ca. 20 µg einzelsträngige JT37/25-DNA
5 nM Gen 49-Deletionsprimer (s. 2.4.)
10 µl TM-Puffer
ad 50 µl

Die Hybridisierung erfolgte für 1h bei 60°C, danach wurde der Ansatz bei Raumtemperatur abgekühlt. Alle weiteren Reaktionen wurden ebenfalls bei Raumtemperatur durchgeführt.

Die Genauigkeit der Hybridisierung wurde überprüft, indem ein Teil des Ansatzes in eine Sequenzierungsreaktion eingesetzt wurde. Bei einer korrekten Hybridisierung des Primers an die Sekundärstruktur sollte ein spezifisches Bandenmuster auf dem Autoradiogramm zu erkennen sein. Eine unkorrekte Hybridisierung des Primers an verschiedenen Regionen der DNA des Klons JT37/25 würde dagegen einen "DNA-Schmier" oder zusätzliche Banden auf dem Autoradiogramm zur Folge haben.

Nach Kontrolle der Bindung des Oligonukleotids an die Sekundärstruktur wurde eine Primer-Verlängerungsreaktion für 10 bzw. 20 Minuten durchgeführt. Die verbliebenen einzelsträngigen DNA-Bereiche des M13-Vektors und des restlichen Inserts wurden durch eine 30-minütige Inkubation mit S1-Nuklease entfernt. Doppelsträngige DNA-Fragmente einer Größe von 500 - 1000 bp (500 bp > Gen 49) wurden nach der Gelelektrophorese aus dem 1,2% Agarosegel isoliert und über stumpfe Enden in einen mit SmaI geschnittenen und dephosphorylierten M13 mp18-Vektor kloniert.

10 Klone wurden isoliert und sequenziert: 4 Klone hatten das Fragment in der Orientierung inseriert, daß der Anfang von Gen 49 sequenziert wurde. Das bedeutet, daß Gen 49 entgegen der Transkriptionsrichtung des lac-Promotors orientiert ist. Bei allen vier Klonen war die Sekundärstruktur Sk-1 deletiert, gleichzeitig war aber bei drei Klonen zusätzlich die Ribosomenbindestelle entfernt. Nur das Insert des Klons JT41/17 enthält Gen 49 ohne Sk-1 mit der ursprünglichen Shine-Dalgarno-Sequenz (s. Tab. 6). Bei den Klonen JT41/3, JT41/8 und JT41/19 ist die Sequenz stromaufwärts des ATG-Initiationskodons in unterschiedlichem Ausmaß deletiert.

Da diese Fragmente aber in die SmaI-Schnittstelle von M13 mp18 kloniert wurden und sich benachbart dieser Schnittstelle die Erkennungssequenzen für die Endonukleasen XbaI (TCTAGA) und Bam HI (GGATCC) befinden, wurde eine "künstliche" Ribosomenbindestelle vor dem ATG-Kodon erzeugt (s. Abb. 17).

Da wahrscheinlich aufgrund von unspezifischer S1-Nuklease-Aktivität DNA-Fragmente unterschiedlicher Länge entstanden, wurde durch diese Methode ein Satz an DNA-Molekülen mit unterschiedlichen Abständen zwischen dem ATG-Kodon und der Ribosomenbindestelle konstruiert. Weil gerade der Abstand ein wichtiger Parameter für die Stärke einer Ribosomenbindestelle darstellt, bietet diese Technik die Möglichkeit, DNA-Fragmente zu erzeugen, welche ein

bestimmtes Gen unter Kontrolle von Shine-Dalgarno-Sequenzen unterschiedlicher Stärke besitzen.

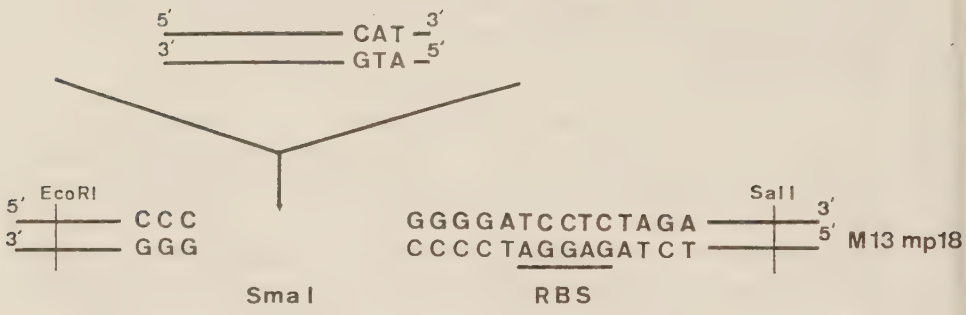


Abb. 17: Konstruktion eines Sets von DNA-Fragmenten mit verschiedenen Ribosomenbindestellen.

Werden DNA-Fragmente mit Gen 49 ohne Ribosomenbindestelle aber mit einem ATG-Initiationskodon und einer unterschiedlichen Anzahl von Nukleotiden stromaufwärts dieses Startkodons in M13 mp18 kloniert, so erhalten alle diejenigen DNA-Fragmente, welche in der hier gezeigten Orientierung einkloniert sind, eine "neue" Shine-Dalgarno-Sequenz. Diese setzt sich aus den Erkennungssequenzen der beiden Restriktionsenzyme XbaI und Bam HI zusammen.

Klon	Ribosomenbindestelle	RB-ATG	Insertgröße
JT41/3	TCTAGAGGATCC↓ATT ATG	7bp	ca. 700bp
JT41/8	TCTAGAGGATCCC↓ATG	5bp	ca.1000bp
JT41/17	GGTAAGGGTTTATT↓ATG	7bp	ca.1000bp
JT41/19	AGAGGATCCCC TTTATT↓ATG	12bp	ca. 300bp

Tab. 6: Ribosomenbindestellen von Gen 49 der Klone JT41/3, JT41/8, JT41/17 und JT41/19.

Die Shine-Dalgarno-Sequenzen und das Initiationskodon sind unterstrichen. Der Pfeil gibt die Übergangsstelle zwischen der M13-und der T4-Sequenz an. Ebenso ist der Abstand zwischen der Ribosomenbindestelle und dem Initiationstriplett sowie die ungefähre Insertgröße angegeben.

Wie zu sehen ist, enthält der Klon JT41/19 nicht das vollständige Gen 49.

Die Klone JT 41/3, JT41/8 JT41/17 und JT41/19 wurden angezüchtet, die doppelsträngige DNA isoliert und das jeweilige Insert mit den Restriktionsendonukleasen Eco RI und SalI herausgeschnitten und über präparative Gele isoliert.

3.4.5. Klonierung von Gen 49 ohne Sekundärstruktur Sk-1 in das pT7-Epressionssystem

Die Insert-DNA der oben genannten Klone wurde über die SalI-Eco RI-Schnittstelle in einen ebenfalls mit diesen beiden Endonukleasen geschnittenen pT7-4-Expressionsvektor ligiert. Die resultierenden Plasmide wurden als pJT7-4/41-3, pJT7-4/41-8, pJT7-4/41-17 und pJT7-4/41-19 bezeichnet.

Mit Ausnahme von pJT7-4/41-19 konnten diese Plasmide, welche Gen 49 ohne die Sekundärstruktur Sk-1 inseriert enthalten, nicht in den Stamm *E. coli* K38/pGP1-2 transformiert werden.

Da das Insert von JT41/19 etwa 300 bp groß ist, kodiert es somit nicht für das vollständige Genprodukt 49. Auf dem Autoradiogramm ist nach radioaktiver Markierung tatsächlich die Überproduktion eines etwa 9 kd großen Proteins festzustellen (s. Abb. 18). Da im Plasmid pT7-4 das ß-Laktamase-Gen in Transkriptionsrichtung des Ø10-Promotors angeordnet ist, wurde in diesen Versuchen auch die ß-Laktamase radioaktiv markiert. Die drei Banden entsprechen dem 29 kd großen Vorläufermolekül und dem prozessierten Protein. Wie zu erkennen ist, wird nahe der Lauffront ein Protein in deutlicher Menge überproduziert.

Warum können Plasmide mit dem vollständigen Gen 49 aber ohne die Sekundärstruktur Sk-1 nicht in *E. coli* K38/pGP1-2 transformiert werden?

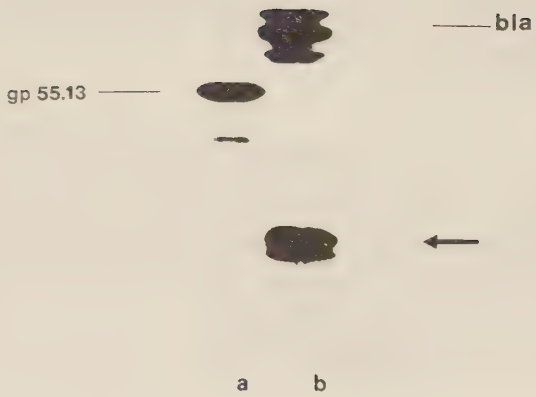


Abb. 18: Induktion des Klons pJT7-4/41-19 und radioaktive Markierung der Proteine.

Die Induktion erfolgte durch Hitzeschock. Als Kontrolle wurde das von dem Klon überproduzierte 20 kd große Genprodukt 55.13 aufgetragen. Die beiden Kanäle zeigen:
a) Klon pJT7-6/49, induziert, mit Rifampicin.
b) Klon pJT7-4/41-19, induziert, mit Rifampicin.
Mit "bla" ist die überexprimierte ß-Laktamase gekennzeichnet, der Pfeil zeigt die Position des von pJT7-4/41-19 synthetisierten Proteins.

Unter der Annahme, daß das Genprodukt 49 auch in relativ geringer Konzentration für die Zellen toxisch ist, kann diese Beobachtung folgendermaßen erklärt werden: Der P_L-Promotor, von dem aus das Gen für die T7-RNA-Polymerase transkribiert wird, ist auch bei 30°C nicht vollständig durch den thermosensitiven cI-Repressor blockiert. Deshalb ist in den Zellen auch im nichtinduzierten Zustand (30°C) eine geringe Menge an T7-RNA-Polymerase vorhanden. Wird das pT7-Plasmid mit Gen 49 in diese Zellen transformiert, so beginnen die wenigen RNA-Polymerase-Moleküle ausgehend von dem Ø10-Promotor mit der Transkription von Gen 49. Da die Sekundärstruktur Sk-1 bei diesen Plasmiden deletiert ist, kann das Genprodukt 49 auch ohne vorherige Induktion in einer für die Zellen letalen Konzentration synthetisiert werden.

Für diese Hypothese spricht, daß Plasmide, welche Gen 49 mit Sk-1 inseriert enthalten, ohne Probleme in K38/pGP1-2 transformiert werden können. Auch hier wird die T7-RNA-Polymerase mRNA ausgehend vom Ø10-Promotor synthetisieren. Diese mRNA kann aber die Sekundärstruktur ausbilden und somit die Translation von Gen 49 weitestgehend verhindern. Ebenso kann das Plasmid pJT7-4/41-19 in K38/pGP1-2 transformiert werden, da es nur, wie in Abb. 19 gezeigt, für ein kleines, vermutlich inaktives aminoterminales Proteinfragment von Genprodukt 49 kodiert.

Um die Hypothese mit experimentellen Beweisen zu unterstützen, wurde je 5µl Ligationsansatz (ca. 0,5 µg DNA) von pJT7-4/41-3 zur Transformation von E. coli K38 ohne pGP1-2 und E. coli K38/pGP1-2 eingesetzt. Der Hitzeschock und die 90-minütige Bebrütung nach der Transformation erfolgten für beide Ansätze bei 30°C. Die Transformationsansätze wurden auf Ampicillin/Kanamycin-Platten (für K38/pGP1-2) bzw. nur auf Ampicillin-Platten (für K38) ausgespatelt. Nach Bebrüten der Platten bei 30°C über Nacht waren bei E. coli K38/pGP1-2 16 Transformanden zu erkennen, bei E. coli K38 ohne pGP1-2 dagegen über 15000! Dieses Ergebnis unterstützt die Hypothese, wonach aufgrund des T7-RNA-Polymerase-Gens in E. coli K38/pGP1-2 Genprodukt 49 auch im uninduzierten Zustand in einer für die Zellen toxischen Konzentration synthetisiert wird.

Die Klonanalyse ergab, daß von den 16 Klonen in K38/pGP1-2 vier ein rekombinantes pT7-Plasmid enthalten. Von den vier Klonen wurde nach Induktion lediglich von pJT7-4/41-3-15M ein Protein von etwa 18 kd in großer Menge synthetisiert. Das Protein ist kleiner als das 20 kd große Genprodukt 55.13 (s. Abb. 19). Zusätzlich zu diesem 18 kd-Protein wird noch ein etwa 10 kd großes Protein synthetisiert. Hierbei handelt es sich möglicherweise um das N-terminale Fragment von Genprodukt 55.13.

Da das 18 kd große Protein in großer Konzentration synthetisiert wird, wurde angenommen, daß der Klon pJT7-4/41-3-15M für ein inaktives Genprodukt 49 kodiert (der Zusatz "M" steht deshalb für "Mutation", siehe auch 3.4.8.).

Von den ca. 15000 Transformanden in E. coli K38 ohne pGP1-2 wurden 20 angezüchtet. Die Klonanalyse ergab, daß alle 20 Plasmide ein inseriertes DNA-Fragment enthalten. Im Vergleich zur Kontrolle (pT7-4 ohne Insert) wurde bei allen rekombinanten Plasmiden nach der Gelelektrophorese ein "DNA-Schmier" festgestellt.

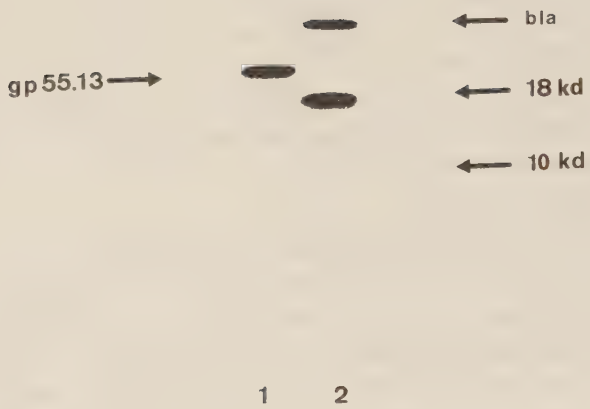
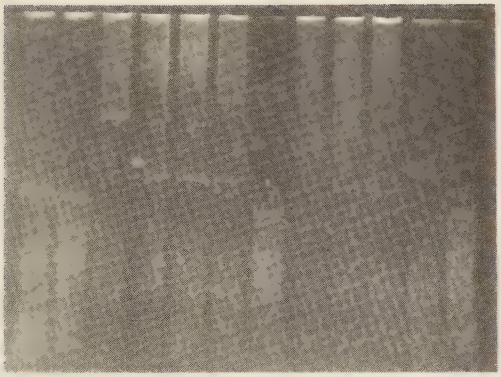


Abb. 19: Autoradiogramm eines 15% Laemmli-Gels.

Die einzelnen Kanäle zeigen:
1) pJT7-6/49, induziert, mit Rifampicin
2) pJT7-4/41-3-15M, induziert, mit Rifampicin

Die DNA-Präparationen erfolgten nach Birnboim und Doly (1979) ohne nachfolgende Reinigung durch eine CSCl-Dichtegradienten-Zentrifugation. Es liegt nahe, daß es sich bei diesem "DNA-Schmier" um Abbauprodukte des E. coli-Chromosoms handelt. Wenn die Vermutung zutreffen sollte, worauf ist dann die Degradierung der E. coli-DNA zurückzuführen?

Untersuchungen ergaben, daß der "DNA-Schmier" nur bei denjenigen Klonen festzustellen ist, welche ein vollständiges Gen 49 ohne Sk-1 tragen. Außerdem erscheint der DNA-Abbau bei Zellen, welche Plasmide mit Gen 49 ohne Sk-1 und mit unterschiedlichen Ribosomenbindestellen enthalten, verschieden stark ausgeprägt (s. Abb. 20). Klone, bei denen die Expression von aktivem Genprodukt 49 durch die Sekundärstruktur Sk-1 inhibiert ist, zeigen dagegen keinen feststellbaren DNA-Abbau; das gilt ebenso für den Klon pJT7-4/41-19, der nur für das aminoterminale Proteinfragment von gp 49 kodiert. Auch der Klon pJT7-4/41-3-15M (s.o.) zeigt keine DNA-Degradierung, was die Vermutung unterstützt, daß dieser Klon für ein inaktives Genprodukt 49 kodiert.



1 2 3 4 5 6 7 8 9 10 11 12

Abb. 20: DNA-Gelelektrophorese von Rohextrakten aus verschiedenen Gen 49 tragenden Klonen.

Die einzelnen Kanäle zeigen:
1-2 : pJT7-4/41-19 (unvollst. Gen 49, -Sk-1, Rb.-ATG: 12)
3-4 : pJT7-4/41-17 (Gen 49, -SK-1, Rb.-ATG: 7)
5-6 : pJT7-4/41-8 (Gen 49, -Sk-1, Rb.-ATG: 5)
7 : pT7-4 ohne Insert
8-10 : pJT7-4/41-3 (Gen 49, -SK-1, Rb.-ATG: 7)
11 : pJT7-5/49 (Gen 49, +Sk-1, Rb.-ATG: 7)
12 : pT7-4 ohne Insert
Hierbei gibt "Rb.-ATG" den Abstand zwischen der Ribosomenbindestelle und dem ATG-Initiationskodon an.

Gen 49 kodiert für eine Endonuklease, welche *in vitro* kruziforme DNA erkennt und schneidet. Eventuell weist der DNA-Abbau darauf hin, daß auf dem *E. coli*-Chromosom palindromische Sequenzen vorkommen, welche von Genprodukt 49 *in vivo* als Substrat erkannt werden.

Wie aber ist zu erklären, daß ein Gen, welches unter der Kontrolle des T7-Promotors steht, auch in Zellen exprimiert wird, denen das Gen für die T7-RNA-Polymerase fehlt? Der T7-Promotor soll nicht von der *E. coli*-RNA-Polymerase erkannt werden (Tabor & Richardson, 1985). Vielleicht initiiert die Wirts-RNA-Polymerase unspezifisch auf dem pT7-Plasmid die Transkription. Da stromaufwärts von Gen 49 auf dem Plasmid keine Terminatoren vorkommen, würde bei einer Transkription von geeigneten Startstellen auf dem pT7-Plasmid

auch Gen 49 in mRNA überschrieben werden. Weil der T7-Promotor Ø10 entgegen der Transkriptionsrichtung ausgehend von Plasmidpromotoren angeordnet ist, kann es sich bei diesen Startstellen nur um sehr schwache Promotoren handeln. Das mag erklären warum gp 49 nur in so geringen Konzentrationen vorkommt, welche die Zellen noch tolerieren.

Bevor die Arbeiten beschrieben werden, die dazu führen, Genprodukt 49 in größeren Mengen zu überproduzieren, sollte festgestellt werden, ob es sich bei dem von dem Klon pJT7-4/41-3-15M kodierten, inaktiven Protein tatsächlich um Genprodukt 49 handelt.

3.4.6. Reinigung und teilweise Sequenzierung des von pJT7-4/41-3-15M kodierten Proteins

Zur Überproduktion des Proteins wurden die Zellen *E. coli* K38/pGP1-2(pJT7-4/41-3-15M) in 2l TBY-Medium angezüchtet (s. 2.13.2.), induziert und geerntet. Der Zellaufschluß erfolgte wie unter 2.12. beschrieben. Die Reinigung des Rohextraktes über eine DNA-Zellulosesäule folgte dem Protokoll von Thomas (1974) und wurde von Frau Aschke-Sonnenborn durchgeführt.

In Zusammenarbeit mit Dr. Meyer, Institut für physiologische Chemie der Ruhr-Universität Bochum, wurde die Sequenz der ersten 10 aminoterminalen Reste des Proteins bestimmt. Die Aminosäuresequenz des NH_2-Terminus ist mit der aus der DNA-Abfolge abgeleiteten Aminosäuresequenz von Genprodukt 49 identisch.

3.4.7. Induktion von Klonen mit Gen 49 ohne Sk-1 in E. coli JM103 durch M13 mGP1-2

Wie kann in Zellen ohne pGP1-2 die Überproduktion von Gen 49 induziert werden? In Abschnitt 2.13. wurde beschrieben, daß die Induktion von Zellen mit dem pT7-Plasmid auch durch Überinfektion mit einem rekombinanten M13-Phagen erreicht werden kann. Dieser Phage, M13 mGP1-2, besitzt das Gen für die T7-RNA-Polymerase unter der Kontrolle des lac-Promotors (s. Abb. 21).

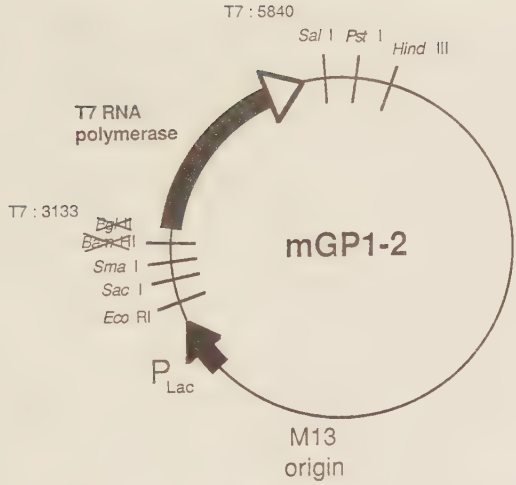


Abb. 21: Das Genom des Phagen M13 mGP1-2.

Das Gen für die T7-RNA-Polymerase wurde mittels BglI-und SalI-Linker in M13 mp8 kloniert. Die Ligation hat den Verlust der BglII-und Bam HI-Schnittstelle zur Folge. Das Gen für die T7-RNA-Polymerase ist unter der Kontrolle des lac-Promotors. Die Gesamtgröße von M13 mGP 1-2 beträgt 9960 bp (S. Tabor, pers. Mitteilung).

Nach Überinfektion und Zugabe von IPTG wird die T7-RNA-Polymerase induziert, und die Transkription von Genen unter der Kontrolle des Ø10 Promotors erfolgt. Grundvoraussetzung ist, daß die Wirtszellen das F-Plasmid besitzen und somit die zur Infektion benötigten F-Pili ausbilden können.

Der von Tabor und Richardson (1985) verwendete E. coli K38-Stamm besitzt das F-Plasmid und ist nicht supprimierend. Das Gen 1 des Phagen T7 ist in M13 mp8 kloniert, der zwei amber-Mutationen in essentiellen Genen aufweist. Deshalb kann der Phage M13 mGP1-2 zwar über die F-Pili die Zellen infizieren und somit zur Induktion der T7-RNA-Polymerase verwendet werden, eine Vermehrung findet aber wegen der fehlenden Suppression nicht statt.

Um zu überprüfen, ob der in unserem Labor vorhandene Stamm E. coli K38 von M13 infiziert werden kann, wurden E. coli JM103-Zellen mit M13 mp18 versetzt und anschließend auf E. coli K38 und E. coli JM103 getitert. Im Gegensatz zu M13 mp8 besitzt M13 mp18 keine amber-Mutation und benötigt deshalb auch keinen supprimierenden Wirt zur Vermehrung.

Während auf E. coli JM103 zahlreiche Plaques zu erkennen waren, fand auf E. coli K38 keine M13-Vermehrung statt. Somit kann dieser Stamm anscheinend nicht von M13-Phagen infiziert werden und ist deshalb nicht für eine Induktion durch M13 mGP1-2 geeignet.

Aus diesem Grund wurden die Gen 49 tragenden Plasmide in E. coli JM103 transformiert. Die Induktion durch M13 mGP1-2 wurde wie unter 2.13.1. beschrieben durchgeführt. Die Autoradiogramme zeigten, daß nach Induktion zwar gp 55.13 und das vermutlich inaktive Genprodukt 49 des Klons pJT7-4/41-3-15M überproduziert werden können, nicht aber aktives Genprodukt 49.

Wie ist dieser Sachverhalt zu erklären? Von Kemper & Garabett (1981) ist für Endonuklease VII eine nukleolytische-Aktivität für einzelsträngige M13 DNA in vitro beschrieben worden. Sollte Endonuklease VII die Aktivität auch in vivo aufweisen, so würde die infizierende einzelsträngige DNA von M13 mGP1-2 zumindest teilweise degradiert werden. Hierdurch wäre die Induktion durch M13 mGP1-2 bei Zellen mit aktivem gp 49 stark abgeschwächt bzw. blockiert.

Sollte diese Annahme zutreffen, so würde sich zur Überproduktion von gp 49 die Aufgabe stellen zu verhindern, daß im uninduzierten Zustand aktive Endonuklease VII in den Zellen synthetisiert werden kann.

Bevor diese Arbeiten begonnen wurden, sollte zuerst durch Aktivitätstests festgestellt werden, ob tatsächlich aktive Endonuklease VII in Gen 49 Klonen nachzuweisen ist.

3.4.8. Endonuklease VII-Aktivitätstests

Diese Experimente wurden in Zusammenarbeit mit Herrn Prof. Dr. B. Kemper und Herrn Dipl. Biol. H. Kosak vom Institut für Genetik der Universität Köln durchgeführt. In Tabelle 7 sind die Ergebnisse der Tests gezeigt.

Tab. 7: Tests auf Endonuklease VII-Aktivität.

Als Kontrollen wurden T4-infizierte Zellen (Positivkontrolle) und Zellen mit dem pT7-6 Plasmid ohne Insert(Negativkontrolle) getestet. Da die Aktivitätsmessungen in mehreren Arbeitsgängen durchgeführt wurden, schwankt der Wert der spezifischen Aktivität (A) zwischen 1,1 und 2,3 Einheiten. Hierbei ist eine Einheit definiert als die Menge an Endonuklease VII, welche 50% von VFS-DNA in Produkte geringeren Molekulargewichts überführt (Kemper & Garabett, 1981).

Die Ribosomenbindestelle des jeweiligen Gen 49 ist ebenfalls abgebildet. Der Quotient $Q=A_x:A_{T4}$ gibt den relativen Anteil an Endonuklease VII-Aktivität des jeweiligen Klons im Vergleich zur Aktivität in T4-infizierten Zellen an. Die Induktion erfolgte durch Infektion mit M13 mGP1-2 bei einer Infektionsmultiplizität von ca. 10-100. Eine Ausnahme bilden die Klone pJT7-5/49 und pJT7-4/41-3-15M. Hier sind die Plasmide in E. coli K38/pGP1-2 transformiert; somit erfolgte die Induktion durch einen Hitzeschock.

	Klon	induziert	Sk-1	Ribosomenbindestelle	spez. Akt (A)	$Q=A_x:A_{T4\text{-inf. Zellen}}$
I				-		
	T4-inf. Zellen			AAGGGTTTATTATG (7 bp)	1,1	1
	pJT7-4/41-8	+	-	AGAGGATCCCATG (5 bp)	0,9	0,82
	pJT7-4/41-17	+	-	AAGGGTTTATTATG (7 bp)	0,8	0,73
	pJT7-4/41-19	+	-	AGAGGATCCCCTTTATTATG (12 bp)	0	0
II						
	T4-inf. Zellen			AAGGGTTTATTATG (7 bp)	2,3	1
	pT7-6	ohne Insert			0	0
	pJT7-6/41-3	-	-	AGAGGATCCATTATG (7 bp)	2,2	0,96
	"	+	-	"	5,7	2,48
	pJT7-5/49	-	+	AAGGGTTTATTATG (7 bp)	0,2	0,09
	"	+	-	"	1,0	0,43
III						
	pJT7-4/41-3-15M	+	-	AGAGGATCCATTATG (7 bp)	0	0

Die Ergebnisse lassen sich wie folgt zusammenfassen:

1. Nur Zellen, die ein Plasmid enthalten, in welchem das vollständige Gen 49 kloniert ist, zeigen Endonuklease VII-Aktivität. So ist bei pJT7-4/41-19, welcher für das aminoterminale Proteinfragment von Endonuklease VII kodiert, keine Enzymaktivität feststellbar.

2. Uninduzierte Zellen mit Gen 49 (± Sk-1) weisen Endonuklease VII Aktivität auf. Bei Klonen mit Gen 49 + Sk-1 (hier: pJT7-5/49) ist eine geringe Aktivität nachweisbar. Nach Induktion durch Hitzeschock steigt die Aktivität etwa auf das Fünffache an. Sie liegt aber immer noch um etwa 50% niedriger als bei T4-infizierten Zellen.

3. Zellen mit Gen 49 ohne Sk-1 weisen im nichtinduzierten Zustand eine 10x höhere Aktivität auf als induzierte Zellen, bei denen aber die Translation von Genprodukt 49 durch die Sekundärstruktur Sk-1 inhibiert werden kann.
4. Die Induktion von pJT7-4/41-3 (Gen 49 ohne Sk-1) durch M13 mGP1-2 erhöht die Enzymaktivität um den Faktor 2,5. Hier liegt eine erste Überproduktion von Genprodukt 49 um den Faktor 2 im Vergleich zu T4-infizierten Zellen vor.
5. In uninduziertem Zustand weisen die drei Klone, welche Gen 49 ohne Sk-1, aber mit verschiedenen Ribosomenbindestellen besitzen, eine etwa gleich große Enzymaktivität auf. Hierbei besitzt der Klon mit der ursprünglichen Ribosomenbindestelle (pJT7-4/41-17) eine geringere Aktivität als die beiden Klone mit neu konstruierten Shine-Dalgarno Sequenzen.
6. Der Klon pJT7-4/41-3-15M kodiert tatsächlich für ein inaktives Genprodukt 49.

Die Messungen der Enzymaktivität unterstützen also die bisher aufgestellten Arbeitshypothesen.

Das Ziel weiterer Arbeiten bestand nun darin, Strategien zu entwickeln, wie ein für die Zellen toxisches Protein in noch größeren Mengen überproduziert werden kann.

3.4.9. Strategien zur Produktionssteigerung von aktiver Endonuklease VII

Wie in Tabelle 7 gezeigt ist, wurde durch Induktion mit M13 mGP1-2 eine Überproduktion von Endonuklease VII um den Faktor 2 im Vergleich zu T4-infizierten Zellen erreicht. Ein Grund für die geringe Induktionseffektivität ist vermutlich das Vorhandensein geringer Mengen dieses Enzyms in den Zellen. Wie können die Schwierigkeiten umgangen werden?

Studier & Moffatt (1986) haben einen rekombinanten Lambda-Phagen konstruiert, der das Gen für die T7-RNA-Polymerase besitzt. Die infizierende doppelsträngige Lambda-DNA wäre kein Substrat für eine eventuell *in vivo* vorhandene einzelstrang-spezifische Nuklease-Aktivität von Genprodukt 49.

Aufgrund der Größe des λ-Genoms ist aber nicht auszuschließen, daß hier palindromische Sequenzen auf der DNA vorkommen, die von Endonuklease VII erkannt und geschnitten werden. Hierdurch würde es ebenfalls zu einer Reduzierung der Induktionseffektivität kommen.

Da anscheinend die Schwierigkeit, Endonuklease VII zu überproduzieren in der Aktivität des Enzyms begründet ist, sollten sich die weiteren Arbeiten darauf konzentrieren, die Endonuklease VII Aktivität in uninduzierten Zellen entscheidend zu reduzieren.

3.4.9.1. Konstruktion eines pT7-Vektors mit einer Struktur zur Transkriptionstermination

Ziel der Arbeiten war es, durch Insertion einer Terminatorstruktur stromaufwärts des T7-Promotors Ø10 von pT7-6 sicherzustellen, daß ohne vorherige Induktion keine Transkription von Gen 49 durch plasmidkodierte Promotorähnliche Strukturen stattfindet.

Als geeignete Insertionsstelle in den Vektor wurde die singuläre PvuII-Schnittstelle ausgewählt (s. Abb. 22).

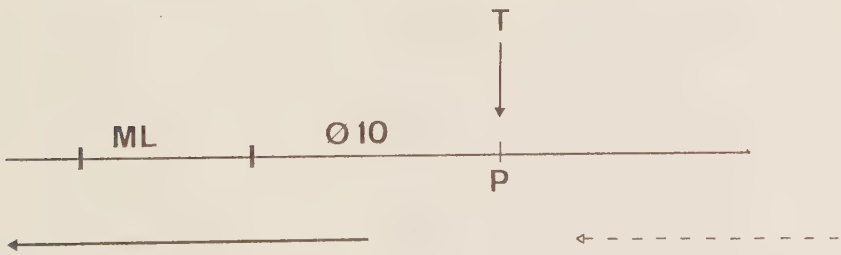


Abb. 22: Schematische Darstellung der Position des Multilinkers, T7-Promotors und der PvuII-Schnittstelle (P) auf dem Plasmid pT7-6.

Der durchgezogene Pfeil gibt die Transkription ausgehend vom T7-Promotor an. Der gestrichelte Pfeil zeigt die potentielle unspezifische Transkription ausgehend von plasmidkodierten Strukturen. Durch Insertion einer geeigneten Terminationsstruktur soll im uninduzierten Zustand die unspezifische Transkription in Richtung einklonierter Gene unterbunden werden.

Der Vektor pT7-6 wurde mit dem Restriktionsenzym PvuII linearisiert und anschließend zur Reduktion der Selbstligationsrate mit alkalischer Phosphatase dephosphoryliert.

Als eine stark terminierend wirkende Sequenz sollte der fd-Terminator verwendet werden. Eine genaue Betrachtung der Sequenz zeigt jedoch, daß direkt stromaufwärts des Terminators eine Promotorstruktur lokalisiert ist. Wird somit das DNA-Fragment in der für die Transkriptionstermination notwendigen Orientierung in pT7-6 kloniert, so kann von dem stromabwärts des Terminators lokalisierten Promotors erneut die Synthese der mRNA in Richtung einklonierter Gene erfolgen. Da der Promotor nicht durch eine geeignete Restriktionsschnittstelle von dem Terminator getrennt werden kann, wurde diese Struktur nicht zur Transkriptionstermination verwendet.

Um zu vermeiden, daß bei anderen möglichen Terminatorstrukturen ebenfalls von eventuell vorhandenen Promotorstrukturen die Transkription in Richtung des einklonierten Gen 49 initiiert werden kann, wurde eine andere Art von Terminatorstruktur ausgewählt: So sollen poly (dT)-Abfolgen eine Termination der Transkription bewirken (R. Eichenlaub, pers. Mitteilung). Da die Transkription in Richtung des Multilinkers von pT7-Plasmiden nur von unkartierten, sehr schwachen promotor-ähnlichen Strukturen erfolgen kann, sollten die poly (dT)-Sequenzen zur Vermeidung der unspezifischen Transkription von einklonierten Genen im uninduzierten Zustand ausreichen.

Als Ausgangsmaterial wurde poly(dA)·poly(dT) mit einer Molekülgröße von etwa 100000 bp von Boehringer Mannheim verwendet. Etwa 30 µg DNA wurden beschallt und Subfragmente einer Größe von etwa 200 - 300 bp isoliert (s.2.11.2.2.). Diese Moleküle wurden über stumpfe Enden in die PvuII-Schnittstelle des Plasmids pT7-6 kloniert. Wird das poly(dA)·poly(dT)-Fragment so in die PvuII-Schnittstelle inseriert, daß das 3´Ende des poly(dT)-Strangs zum Multilinker orientiert ist, sollte es zur Transkriptionstermination kommen.

Eine Klonanalyse ergab, daß die Plasmide pJT7-6TpAT6, pJT7-6TpAT8, pJT7-6T pAT18 und pJT7-6TpAT19 ein poly(dA)·poly(dT)-Fragment inseriert enthalten (das "T" steht für "Terminator", "pAT" bedeutet "poly(dA)·poly(dT)").

Die vier Plasmide wurden isoliert und mit den Restriktionsendonukleasen SalI und Eco RI inkubiert. In der gleichen Weise wurde das DNA-Insert aus M13 JT41/8 isoliert. Dieses Fragment beherbergt Gen 49 ohne Sk-1 sowie den offenen Leserahmen 55.13. Je 0,3 µg Insert-DNA wurden in eines der 4 poly(dA)·poly(dT) enthaltenden Plasmide ligiert und die Ligationsansätze in E. coli

JM103 transformiert. Da die Plasmide vorher mit Eco RI und SalI geschnitten wurden, erfolgte die Insertion von Gen 49 in Transkriptionsrichtung des Ø10-Promotors.

Die Klonanalyse zeigte folgendes Ergebnis: Alle rekombinanten Plasmide, welche sich von den ursprünglichen Vektoren pJT7-6TpAT6, pJT7-6TpAT18 und pJT7-6TpAT19 ableiten, zeigen mit wenigen Ausnahmen Anzeichen von Endonuklease VII-Aktivität in Form von DNA-Degradierung nach Gelelektrophorese der Rohextrakte. Dagegen ist bei den Gen 49 tragenden Plasmiden, die sich von dem Vektor pJT7-6TpAT8 ableiten, kein DNA-Abbau festzustellen. Das deutet möglicherweise auf eine Blockierung der von Plasmidstrukturen ausgehenden Transkription durch Einbau des poly(dA)·poly(dT)-Fragments in der terminierenden Orientierung hin.

Der Ligationsansatz, bei dem das Insert von JT41/8 in den Vektor pJT7-6TpAT8 ligiert wird, wurde in E. coli K38/pGP1-2 transformiert. Die rekombinanten Klone wurden induziert und die überproduzierten Proteine mit ^{35}S-Methionin markiert. Das Ergebnis der Versuche ist in Abbildung 23 dargestellt.

1 2 3

Abb. 23: Überexpression von Endonuklease VII.

1 & 2: pJT7-6TpAT8/41-8-1 und pJT7-6TpAT8/41-8-4, induziert
3: pJT7-4/41-3-15M, induziert

Es werden zwei Proteine überproduziert. Das kleinere Protein zeigt ein ähnliches Molekulargewicht wie das von dem pJT7-4/41-3-15M synthetisierte inaktive Genprodukt 49. Bei dem größeren Protein sollte es sich um Genprodukt 55.13 handeln, da das DNA-Insert aus M13 JT41/8 ca. 1000 bp groß ist und somit orf 55.13 größtenteils oder sogar vollständig enthalten sollte. Weiterführende Experimente ergaben jedoch, daß das von pJT7-6TpAT8/41-8-1 bzw. pJT7-6TpAT8/41-8-4 synthetisierte Protein etwa 3 kd größer ist, als das von pJT7-6/orf 55.13 produzierte Genprodukt. Eine Erklärung hierfür ist, daß das Insert aus M13 JT41/8 nicht den vollständigen Leserahmen 55.13 beinhaltet und es durch einen Rasterschub im Leserahmen zu einem größeren "Hybrid-Gen" aus orf 55.13 und einer Nukleotidsequenz des pT7-Plasmids kommt.

Ist das überproduzierte 18 kd große Protein der rekombinanten pJT7-6TpAT8-Klone aktive Endonuklease VII? Zur Aufklärung dieser Frage wurden aus je 1l Kulturmedium die induzierten Zellen geerntet und die Rohextrakte in Zusammenarbeit mit Dr. B. Kemper und Dipl. Biol. H. Kosak auf Endonuklease VII-Aktivität getestet (s. Tab. 8).

Zellen	spez Akt.(A)	$Q=A_x : A_{T4\text{-inf. Zellen}}$
T4-inf. Zellen	25u/µg Prot.	1
pJT7-6TpAT8/41-8(1)	378u/µg Prot.	15,1
pJT7-6TpAT8/41-8(4)	674u/µg Prot.	27,0

Tab. 8: Endonuklease VII-Aktivitätstests.

Wie zu sehen ist, wird aktive Endonuklease VII von den Klonen mit einem poly(dA)·poly(dT)-Terminator 15 bzw. 27 mal im Vergleich zu T4-infizierten Zellen überproduziert. Zusätzliche Versuche von H. Kosak zeigten, daß im uninduzierten Zustand weniger als ein Hundertstel an Endonuklease VII-Aktivität gemessen werden kann als nach Induktion durch Hitzeschock. Eine Reinigung der klonierten aktiven Endonuklease entsprechend dem unter 2.14 beschriebenen Protokoll, sowie die weiteren biochemischen Untersuchungen des Enzyms werden z. Zt. in Zusammenarbeit mit Dr. B. Kemper durchgeführt.

3.5. Sequenzierung von Gen 49 sowie anliegender Regionen bei den Bakteriophagen T2 und T6

Die geradzahligen Bakteriophagen T2, T4 und T6 sind miteinander nahe verwandt. Große Genomabschnitte weisen eine ausgeprägte DNA-Homologie auf (Kim & Davidson, 1974). So sind zum Beispiel die Nukleotidsequenzen für die alpha-Glukosyltransferasen von T2, T4 und T6 sehr ähnlich. Unterschiede treten bevorzugt in der dritten Position der Tripletts auf ("wobble-position") und haben häufig keine Auswirkung auf die Aminosäuresequenz des Enzyms (Gram & Rüger, 1986).

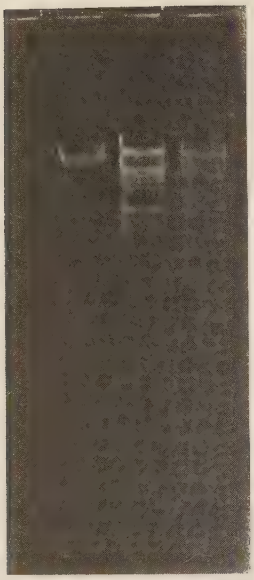
Für die Phagen T2 und T6 ist bisher nichts über eine möglicherweise vorkommende Endonuklease VII Aktivität-bekannt. Aus diesem Grund sollte festgestellt werden, ob hier eine dem Gen 49 von T4 homologe Struktur vorhanden ist.

Hierzu wurde die Hydroxymethylcytosin-haltige DNA von T2 und T6 sowie die dC-DNA von T4 mit der Restriktionsendonuklease Eco RV hydrolysiert (Eco RV ist eines der wenigen Enzyme, die auch HMC-DNA schneiden können). Die drei Restriktionsmuster unterscheiden sich deutlich voneinander (s. Abb. 24A).

Die DNA wurde, wie unter 2.10.3. beschrieben, auf Nylonfilter übertragen und die partiell einzelsträngige, radioaktiv markierte DNA (s. 2.10.1.) des Klons M13 JT37/21 anhybridisiert. Dieser Klon trägt das 1,6 kb große HpaII-PvuII-Fragment, auf dem sowohl Gen 49 als auch orf 55.13 lokalisiert sind. Das Ergebnis der Hybridisierung ist der Abbildung 24B zu entnehmen. Wie zu sehen ist, findet sowohl bei T4, als auch bei T2 und T6 eine spezifische Hybridisierung statt, was auf ein Vorhandensein von orf 49.1, Gen 49 bzw. orf 55.13-Äquivalenten bei T2-und T6-Phagen schließen läßt.

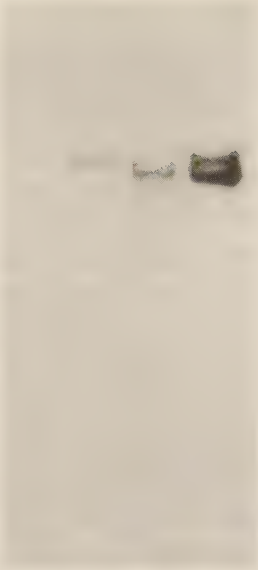
Diese Versuche bildeten die Grundlage zur Sequenzierung der Gen 49-Äquivalente von T2 und T6.

In unserer Arbeitsgruppe ist eine T2-und T6-Klonbank angelegt worden (Gram und Rüger, 1986). Hierzu wurde die T2-bzw. T6-DNA mit Ultraschall zerkleinert und DNA-Fragmente einer Größe zwischen 1 kb und 2 kb nach der Elektrophorese aus dem Agarosegel isoliert. Diese Subfragmente wurden in die HincII-Schnittstelle von M13 mp18 kloniert. Die beiden Klonbanken von T2-und T6-DNA enthalten ca. 8000 bzw. 5000 unterschiedliche Klone.



1 2 3

A



1 2 3

B

Abb. 24: Untersuchungen an der DNA von T2, T4 und T6.

A: Gelelektrophoretische Auftrennung von:
1) T2 HMC-DNA x Eco RV
2) T4 dC-DNA x Eco RV
3) T6 HMC-DNA x Eco RV
Es wurde ein 1,2% Agarosegel verwendet.

B: Autoradiogramm der Hybridisierung von radioaktiv markierter M13 JT 37/21-DNA an die oben gezeigten Restriktionsmuster. Die Reihenfolge ist wie unter A gezeigt.

Als "Sonde" zum Auffinden von Gen 49 homologen T2-und T6-Inserts aus den Klonbanken wurden die rekombinanten M13-Klone JT37/21 und JT37/25 eingesetzt (beide Klone enthalten das 1,6 kb große HpaII-PvuII-Fragment in entgegengesetzter Orientierung).

Die Phagensuspensionen der Klonbanken wurden so verdünnt, daß etwa 5000 Plaques auf einer 82 mm großen Agarplatte geformt wurden. Die Plaquehybridisierung (s. 2.10.2.) wurde mit der radioaktiv markierten DNA der Klone M13 JT37/21 und M13 JT37/25 durchgeführt.

Die Autoradiogramme wurden mit den Agarplatten verglichen, die "positiven" Plaques ausgestochen und die Phagen in TBY-Medium überführt. Die Phagen wurden erneut ausgespatelt (ca. 50 - 200 pro Platte), die DNA auf Nylonfilter übertragen und die jeweilige radioaktive Probe anhybridisiert. Die Schwärzungen auf dem Autoradiogramm konnten eindeutig den entsprechenden Plaques zugeordnet werden. Die DNA dieser Klone wurde sequenziert und die Nukleotidabfolgen anschließend mit Computerprogrammen analysiert. Die Auswertung ergab, daß sowohl orf 49.1, Gen 49 als auch orf 55.13 bei T2 und T6 vorkommen. Da aber beide Nukleotidabfolgen nicht vollständig sequenziert werden konnten, wurden folgende Strategien zur Vervollständigung der beiden Nukleotidsequenzen verfolgt:

1) Durch die Auftrennung der Sequenzierungsprodukte geeigneter Klone auf speziellen langlaufenden 6% Polyacrylamid/Harnstoffgelen ohne Gradient wurde etwa bis zur 340. Base eines Inserts gelesen. Hierdurch konnten mehrere Sequenzlücken geschlossen werden.

2) In die DNA von Klonen mit sehr großen Inserts, bei denen die Sequenzierung der ersten 300 Basenpaare nicht genügte, wurden gezielt Deletionen in die inserierte DNA durch Verdauung mit geeigneten Restriktionsendonukleasen eingeführt. Somit war es möglich, bei einer erneuten Sequenzierung des verkleinerten Inserts weitere Nukleotidabfolgen zu ermitteln.

Zur Sequenzierung der 2067 bp großen DNA-Region von T2 waren 33 Klone notwendig. Für die Erstellung des 1509 bp großen DNA-Abschnitts des Bakteriophagen T6 wurden 35 Klone sequenziert.

3.5.1. Orf 49.1, Gen 49 und orf 55.13 der Bakteriophagen T2, T4 und T6 sind sehr homolog

Sowohl orf 49.1 als auch Gen 49 sind bei allen drei Phagentypen gleich groß. Es gibt keine intragenischen Insertionen oder Deletionen. Der Homologiegrad ist sehr hoch.

Nachfolgend sollen die einzelnen Genbereiche beschrieben werden.

Orf 49.1 bei T2, T4 und T6

Alle drei Genstrukturen besitzen die gleiche Ribosomenbindestelle. Der stromaufwärts gelegene frühe Promotor wurde nur bei T4 vollständig sequenziert. Bei T6 ist aber die -10 Box bestimmt worden. Sie stimmt mit der bei T4 festgestellten Basenabfolge überein.

Insgesamt sind 13 Tripletts von den 161 Kodons bei T2, T4 und T6 unterschiedlich. Diese Tripletts, sowie die daraus abgeleiteten Aminosäuren sind in Tabelle 9 gezeigt.

	Tripletts			Aminosäuren		
Pos.	T2	T4	T6	T2	T4	T6
6	AAG	AAA	AAA	Lys	Lys	Lys
17	CTA	CTG	CTG	Leu	Leu	Leu
31	AAA	AAA	AAG	Lys	Lys	Lys
36	GCA	GTA	GCA	Ala	Val	Ala
44	GAA	GAG	GAA	Glu	Glu	Glu
46	GAT	GAC	GAC	Asp	Asp	Asp
68	AAG	ACG	ACG	Lys	Thr	Thr
69	TAC	TAT	TAT	Tyr	Tyr	Tyr
78	GTC	GTA	GTC	Val	Val	Val
129	GCC	GCC	GCT	Ala	Ala	Ala
146	CGA	CGA	CAA	Arg	Arg	Gln
147	ATG	ATG	AAG	Met	Met	Lys
149	CAG	CAG	CGG	Gln	Gln	Arg

Tab. 9: Orf 49.1 von T2, T4 und T6.

Gezeigt werden die Positionen der unterschiedlichen Tripletts, deren Sequenzen, sowie die von den Kodons abgeleiteten Aminosäureabfolgen bei T2, T4 und T6.

Die meisten Basenaustausche finden sich in der "wobble" Position eines Tripletts. Die Basenaustausche führen nur in fünf Fällen zu einem Aminosäureaustausch.

Gen 49 von T2, T4 und T6

Alle drei Gene besitzen die für die Regulation der Genexpression notwendige Sekundärstruktur Sk-1 mit dem späten Promotor P_{1-1}. Ebenso ist in dem für den Carboxylterminus kodierenden Bereich ein potentieller mittlerer Promotor lokalisiert.

Von den 157 Kodons unterscheiden sich 13: jeweils acht Abweichungen treten bei Vergleich von T2 und T4, bzw. T2 und T6 auf, zwischen den T4- und T6-Klonen gibt es dagegen 12 unterschiedliche Tripletts. Auch hier werden Basenaustausche besonders in der dritten Position beobachtet (s. Tab. 10).

	Tripletts			Aminosäuren		
Pos.	T2	T4	T6	T2	T4	T6
13	CAA	CAC	CAA	Gln	Gln	Gln
17	AAT	GAT	AAT	Asn	Asp	Asn
38	CAC	CAC	CAT	His	His	His
39	CTC	CTC	CTT	Leu	Leu	Leu
40	GAC	GAC	GAT	Asp	Asp	Asp
53	GTG	GTG	GTA	Val	Val	Val
60	CTC	CTA	CTC	Leu	Leu	Leu
61	TGC	TGC	TGT	Cys	Cys	Cys
63	GCA	GCT	GCT	Ala	Ala	Ala
70	CAC	CAT	CAC	His	His	His
81	GGG	GGT	GGG	Gly	Gly	Gly
94	TAC	TAT	TAC	Tyr	Tyr	Tyr
143	CAG	CAA	CAG	Gln	Gln	Gln

Tab. 10: Gen 49 von T2, T4 und T6.

Gezeigt sind die unterschiedlichen Tripletts sowie die hiervon abgeleiteten Aminosäuren.

Der Genbereich um orf 55.13 bei T2, T4 und T6

Wie in Abschnitt 1.5 beschrieben wurde, haben Gott et al. (1986) auf dem Restriktionsfragment XbaI-10 durch radioaktive Endmarkierung von deproteinisierter T4-RNA mit α-^{32}P-dGTP ein Intron der Gruppe I entdeckt.

Im Rahmen dieser Dissertation wurde in Zusammenarbeit mit den Labors von Dr. Shub und Dr. Belfort in Albany (USA) durch Sequenzierung der reifen mRNA die exakte Spleißstelle des Introns bestimmt. Zu diesem Zweck wurde die zum damaligen Zeitpunkt unveröffentlichte Sequenz des XbaI-10-Fragments zur Verfügung gestellt, um die zur Sequenzierung der gespleißten mRNA notwendigen Primer synthetisieren zu können.

Darüberhinaus wurden von mir durch Sequenzierung der T2- und T6-DNA um orf 55.13 diese Ergebnisse weiter abgesichert. Es zeigte sich, daß bei T2 und T6 zwar orf 55.13 und orf 55.11 vorkommen, nicht aber ein etwa 1000 bp großes DNA-Fragment, auf dem der 774 bp große orf 55.12 lokalisiert ist. Weiterhin sind bei T2 und T6 die beiden offenen Leserahmen 55.13 und 55.11 zu einer Genstruktur fusioniert.

Hierbei erfolgte die Sequenzierung der DNA-Abfolge von T6 im Bereich der bei T4 vorkommenden Spleißstelle nicht durch Analyse der T6-Klone, sondern durch Hybridisierung eines geeigneten Primers (MB47, s. 2.4.) an die gesamte T6-DNA und anschließender Primer-Verlängerung mit Didesoxynukleotiden.

Die Vermutung, daß es sich bei orf 55.13/55.11 bei T4 um das von Gott et al. (1986) lokalisierte Mosaikgen handelt, wurde durch RNA-Sequenzierungen bestätigt (Shub et al., 1988). Die beiden Exonstrukturen werden bei T4 durch ein 1033 bp großes Intron unterbrochen. Auf dem Intron befindet sich der offene Leserahmen 55.12. Das T4-Mosaikgen wurde von Shub et al. (1988) vorläufig "sunY" ("split gene, unknown function, why") genannt. In Abb. 25 ist die Spleißsequenz des sunY Introns dargestellt.

Ein Vergleich der Sequenz der gespleißten T4-mRNA mit der T2- und T6-DNA um die Spleißstelle ist in Abb. 26 gezeigt. Da die Sequenzen von T2 und T6 in dieser Region identisch sind, werden sie im weiteren Verlauf der Arbeit als T2/T6-DNA bezeichnet.

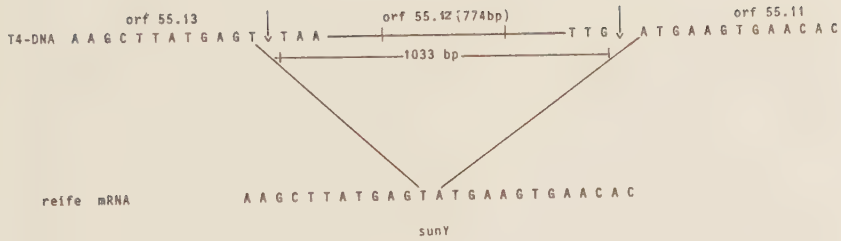


Abb. 25: Darstellung der Spleißsequenz des sunY-Introns.

Gezeigt ist die von der mRNA abgeleitete DNA-Abfolge aus dieser Region. Die Pfeile geben die genaue Spleißstelle an.

Spleißstelle

	gln	ala	tyr	glu	tyr	glu	val	asn	thr	leu	phe	ser					
T4:	CAA	GCT	TAT	GAG	TAT	GAA	GTG	AAC	ACG	TTA	TTC	AGT					
	gln	ala	tyr	glu	tyr	glu	val	asn	thr	leu	his	ser					
T2/T6 :	CAA	GCT	TAT	GAA	TAT	GAA	GTA	AAT	ACT	CTC	TTT	AGT					

Abb. 26: Vergleich der Sequenzen der reifen T4-mRNA mit der T2/T6-DNA um die Spleißstelle des sunY-Introns.

Gezeigt ist die von der gespleißten T4-mRNA abgeleitete DNA-Sequenz und die Nukleotidabfolge der T2-und T6-DNA (T2/T6). Die Spleißstelle bei T4 ist durch einen Pfeil gekennzeichnet. Von T4 abweichende Nukleotide und Aminosäuren sind bei der Sequenz von T2/T6 unterstrichen.

4. DISKUSSION

4.1. Klonierung und Sequenzierung der 9,1 kb-Region des Bakteriophagen T4

Zur Bestimmung der Nukleotidabfolge wurden die Restriktionsfragmente XbaI-18 und XbaI-19 vollständig, und die Fragmente XbaI-10 und XbaI-14 teilweise sequenziert. Die anschließende gezielte Klonierung und Sequenzierung von HindIII, Eco RI und RsaI-Restriktionsfragmenten gewährleistete die einwandfreie Verknüpfung der vier Teilsequenzen an den XbaI-Schnittstellen.

Ein Vergleich mit der von Le Master (1986) publizierten 540 bp großen Sequenz um Gen nrdC (Position 109-649 bp) zeigt eine vollständige Übereinstimmung mit der in dieser Arbeit bestimmten Nukleotidabfolge. Ebenso ist die Gen 49 Basenabfolge mit der von Barth et al. erstellten Sequenz identisch (G. Mosig, pers. Mitteilung).

Von den 9100 bp sind etwa 600 bp in der von Gram & Rüger (1985) erstellten Nukleotidsequenz enthalten. Die Basenabfolgen stimmen überein. Deshalb war es nicht notwendig, in diesem Bereich die Sequenzen beider DNA-Stränge zu bestimmen. Weiterhin wurden folgende Nukleotidabfolgen nur durch Sequenzierung in einer Orientierung bestimmt:

Position		Größe
91 - 246	:	155 bp
696 - 745	:	49 bp
1131 - 1141	:	10 bp
1355 - 1482	:	127 bp
5881 - 5896	:	15 bp
6718 - 6763	:	45 bp
7142 - 7149	:	7 bp
7769 - 7823	:	54 bp
8365 - 8434	:	69 bp

Da die erste in dieser Arbeit aufgeführte Genstruktur (nrdC) erst bei Position 249 kartiert, wurde davon abgesehen, die Nukleotidsequenz von 91-246 bp in beiden Orientierungen zu bestimmen.

Frühe T4-Promotoren können nicht in der Orientierung in M13-Vektoren kloniert werden, die der Transkriptionsrichtung von M13-Promotoren aus entgegengesetzt ist. Hierfür ist möglicherweise eine Behinderung der M13-Transkription durch die von T4-Promotoren ausgehende entgegengesetzte mRNA-Synthese verantwortlich. Tatsächlich konnten die Nukleotidabfolgen der drei frühen Promotoren P_{e-1} bis P_{e-3} durch Sequenzierung der einzelsträngigen M13-DNA nur in einer Orientierung analysiert werden. Die Basensequenz des Gegenstrangs wurde durch DNA-Doppelstrang-Sequenzierung eines rekombinanten M13-Vektors bestimmt, welcher zur Klonierung starker, früher T4-Promotoren konstruiert wurde (Liebig, 1987).

Für die 9,1 kb-Sequenz wurde eine Restriktionskarte erstellt (Staden, 1980). Unter Berücksichtigung einer Fehlerspanne von ± 200 bp bei der Kartierung ergibt sich in der Regel eine gute Übereinstimmung der von der Nukleotidsequenz abgeleiteten Schnittpositionen mit den kartierten Schnittstellen. Eine größere Abweichung findet sich nur bei den Kartenpositionen für die Schnittstellen der Restriktionsendonuklease HindIII. Während O´Farrel et al. (1980) in dieser Region vier Erkennungssequenzen kartiert haben, lassen sich aus der Sequenz elf Schnittstellen ableiten. Diese Abweichungen sind aber in Übereinstimmung mit früheren Erfahrungen, wonach die über eine zweidimensionale Birnstiel-Technik erstellte Karte für die Restriktionsendonuklease HindIII ungenau ist. Außerdem wurde eine zusätzliche Eco RI-Schnittstelle ermittelt. Bei diesem 93 bp großen Restriktionsfragment handelt es sich wahrscheinlich um die Bande Eco RI-47.

4.2. Offene Leserahmen auf der 9,1 kb-Sequenz

Eine Analyse der 9,1 kb großen Sequenz ergab, daß neben den drei in dieser Region kartierten Genen noch mindestens 16 weitere Genstrukturen lokalisiert sind.

Arbeiten von Shub et al. (1988) zeigen, daß es sich bei drei dieser offenen Leserahmen (orf 55.13, orf 55.12 und orf 55.11) um das dritte bekannte T4-Mosaikgen sunY handelt. Die Resultate für dieses Mosaikgen werden getrennt in Abschnitt 4.2.1. diskutiert.

Die potentiellen Gene besitzen Kodierungskapazitäten für Proteine zwischen 5140 d und 67874 d. Die Funktionen der möglichen Genprodukte sind weitgehend unbekannt. Nur für orf 49.1 wurde in Zusammenarbeit mit dem Labor von Dr. L. Simon gezeigt, daß es für einen Proteinase-Inhibitor kodiert (Skorupski et al., 1988). Eine Diskussion der Ergebnisse für orf 49.1 erfolgt in Kapitel 4.2.2.

Die Genstrukturen innerhalb der 9,1-kb Region sind in der Regel eng benachbart angeordnet. Überlappungen zwischen einzelnen Genen kommen häufig in Form der ATGA-Elemente vor (s. 3.2.2.) und kartieren zwischen den Genen nrdC/49.3, 49.3/49.2, Gen 49/55.13 sowie zwischen 55.11/55.10. Möglicherweise deuten diese Überlappungen auf eine enge translationale Kopplung der Genstrukturen hin. Dabei würde das Ribosom an dem Stopkodon nicht von der mRNA dissoziieren, sondern durch einen um eine Base nach "rückwärts" gerichteten Translokationsschritt die Proteinbiosynthese am nachfolgenden Leserahmen beginnen (Schumperli et al., 1982). Die zu dem Starttriplett gehörende Ribosomenbindestelle liegt hierbei in der kodierenden Sequenz des vorhergehenden Leserahmens; sie unterscheiden sich in ihrer Homologie zum 3´Ende der 16S-RNA voneinander. Während zum Beispiel orf 49.2 eine ausgeprägte Ribosomenbindestelle besitzt, weist orf 55.13 nur eine relativ schwache Komplementarität zum 3´Ende der 16S-RNA auf (s. Tab. 1). Wie aber die in 3.4.3. beschriebenen Versuche zur Überexpression von Gen 49 zeigen, wird Genprodukt 55.13 in einer überraschend hohen Konzentration von einem rekombinanten pT7-Expressionsvektor synthetisiert (über 30% des totalen Zellproteins). Dieser Sachverhalt wird in Abschnitt 4.4. diskutiert.

ATGA-Elemente wurden bei einer Reihe von T4-Genen gefunden, so z.B. zwischen orf 47.1 und Gen 47 (Gram & Rüger, 1985), zwischen Gen 33 und Gen 59 (Hahn, 1986) sowie zwischen den Genen frd und td (Purohit & Matthews, 1984). Diese Strukturen sind aber nicht auf T4 beschränkt. So kommen ATGA-Elemente auch zwischen den Genen O und P des Phagen Lambda (Schwarz et al., 1978) und im Chloroplasten-Genom von Zea mays vor (Krebbers et al., 1982).

Wo zwischen einzelnen Leserahmen nichtkodierende Abschnitte vorkommen, sind in der Regel Strukturen für die Transkriptionsinitiation und Termination angeordnet (s. 4.3.). Zwei große intergenische Strukturen sind Bereiche des sunY-Introns (s. 4.2.1.).

Insgesamt besitzen 90,1% der Nukleotidsequenz aus der Region von 48,266 kb bis 39,166 kb Kodierungskapazität für Proteine. Dieser Anteil entspricht in etwa denen für andere große sequenzierte T4-Genomabschnitte. So sind z.B. von den 7084 bp zwischen den Kartenpositionen 39,4 kb und 32,3 kb 91,8% für Proteine kodierend (Gram & Rüger, 1985); von den 7200 bp aus den Positionen 65,5 kb bis 58,3 kb können 93,0% Genstrukturen zugeordnet werden (Valerie et al., 1986).

In Tabelle 2 sind die aus der Nukleotidsequenz abgeleiteten Proteine charakterisiert. Die Polaritätsindizes liegen zwischen 36 und 62. Membranproteine besitzen in der Regel Indizes unter 40 (Capaldi & Vanderkooi, 1972). Es ist deshalb zu vermuten, daß die von orf 55.8 und orf 55.1 kodierten Proteine (Polaritätsindizes von 39 bzw. 36) möglicherweise membrangebunden sind. Die Polaritätsindizes für gp 49 von 54 und gp 55 von 55 lassen dagegen eher auf lösliche Proteine schließen.

Ein Vergleich der Genanfänge und der daraus abgeleiteten Aminosäuresequenzen zeigt eine Homologie der N-terminalen Bereiche von gp 55.10 und gp 55.9 (s. Abb. 27).

	M	N	Y	D	R	I	Y	S
orf 55.10:	ATG	AAT	TAT	GAC	AGA	ATG	TAT	TCA

	M	N	Y	D	R	I	Y	P
orf 55.9 :	ATG	AAT	TAT	GAT	AGA	ATT	TAT	CCA

Abb. 27: Genanfänge von orf 55.10 und orf 55.9 sowie die daraus abgeleiteten Aminosäuresequenzen.

Die jeweiligen identischen Nukleotide und Aminosäuren sind unterstrichen.

Der Kodongebrauch in T4-Genen unterscheidet sich wesentlich von dem in stark exprimierten E. coli-Genen. So kodiert T4 für acht eigene tRNA-Moleküle (Fukada & Abelson, 1980). Es wird vermutet, daß bei T4-Genen dasjenige Kodon für eine Aminosäure bevorzugt wird, für das eine eigene tRNA-Spezies existiert. Ein Vergleich des Kodongebrauchs von stark exprimierten E. coli - Genen mit dem Kodongebrauch von den in der 9,1 kb-Region angeordneten 19 Genstrukturen ist in Tabelle 12 im Anhang dargestellt.

Die größten Unterschiede ergeben sich erwartungsgemäß bei den Kodons, die komplementär zu den von T4 kodierten Antikodons sind. Weiterhin wird bei T4 im Gegensatz zu E. coli z.B. für Tyrosin eher das Kodon TAT als TAC benutzt. In der Regel werden bei T4-Genen Tripletts mit A und T in der dritten Position häufiger benutzt als solche mit C und G am Ende.

In Tabelle 13 (s. Anhang) sind die Kodonhäufigkeiten für die Gene nrdC, pinA, 49, sunY, 55 und den Intron-kodierten Leserahmen 55.12 dargestellt. Wie zu erkennen ist, unterscheiden sich einige Gene in ihrem Kodongebrauch voneinander. So wird z. B. bei Gen nrdC das Kodon TCA, welches für Serin kodiert und komplementär zu dem Antikodon einer von T4 kodierten tRNA ist, nicht verwendet. Das gilt ebenso für das Triplett CCA, welches für Prolin kodiert.

4.2.1. Das Mosaikgen sunY

Nach den Genen td und nrdB ist sunY das dritte bekannte Mosaikgen von T4. Da alle drei Intron-RNA-Moleküle Guanosin an das 3´Ende binden und anschließend unter Abspaltung eines Oligonukleotids einschließlich des Guanosinrestes zirkularisieren (Ehrenmann et al., 1986, Gott et al., 1986), werden sie den Introns der Gruppe I zugeordnet (Cech, 1986).

Das Gen sunY kodiert für ein 605 Aminosäuren großes Protein, welches ein Molekulargewicht von 67879 d aufweist. Die Nukleotidsequenz des "reifen" Gens und die abgeleitete Aminosäureabfolge finden sich in Abbildung 35 (s. Anhang). Der Polaritätsindex von 47 läßt wahrscheinlich auf ein lösliches Protein schließen. 14,0% des Proteins sind positiv und 11,9% negativ geladen.

Während die Genprodukte von nrdB und td am Nukleotidstoffwechsel beteiligt sind, ist über die Funktion von gp sunY nichts bekannt. Es gibt aber einen Hinweis darauf, daß es sich bei diesem Protein möglicherweise um ein DNA-bindendes Metalloprotein handelt; und zwar ist bei Position 543 - 564 der Aminosäuresequenz eine Struktur lokalisiert, die häufig bei metallatombindenden Proteinen vorkommt ("Zinkfinger"), siehe Abb. 28. Diese Metallbindestelle ist besonders für DNA-Bindeproteine charakteristisch.

Anstelle eines Cysteins kann bei diesen Proteinen auch ein Histidinrest vorkommen. Es läßt sich folgende Konsensussequenz formulieren:

$$C - A_{2-4} - X - A_n - X - A_{2-4} - X \qquad \text{(mod. nach Berg, 1986)}$$

Hierbei steht C für Cystein und X für ein Cystein oder Histidin. A bedeutet eine beliebige Aminosäure, der Abstand zwischen jeweils zwei Cystein- oder Histidinresten beträgt in der Regel 2-4 Aminosäuren. Die Anzahl an Aminosäuren (A_n) zwischen den beiden Cystein/Histidin-Paaren scheint relativ variabel zu sein. Es kommen normalerweise Abstände zwischen 4 und 16 Resten vor (Berg, 1986).

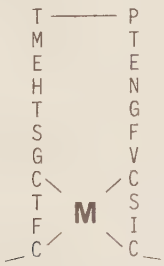


Abb. 28: Potentielle metallbindende Domäne von Genprodukt sunY aus den Positionen 543 - 564.

Der eukaryontische Transkriptionsfaktor TFIIIA von Xenopus laevis, welcher an 5S RNA-Gene und deren Transkripten bindet, besitzt neun Einheiten folgender Sequenz: Cys-A_{2-5}-Cys-A_{12}-His-A_{2-3}-His.

Miller et al. (1985) vermuten, daß diese Abschnitte Zn^{++}-bindende Domänen darstellen, welche mit Nukleinsäuren über hydrophile Reste interagieren können.

Ähnliche, möglicherweise Zink-bindende Strukturen finden sich ebenso bei Endonuklease VII von T4 (s. 4.2.4.), bei der T4-DNA-Primase (Berg, 1986), bei dem Helix-destabilisierenden Genprodukt 32 von T4 (Giedroc et al., 1986), bei dem menschlichen Glucocorticoid-Rezeptor (Weinberger et al., 1985) und dem menschlichen Östrogen-Rezeptor (Greene et al., 1986). Auch das 97 Aminosäuren große, von dem nrdB-Intron kodierte Protein besitzt eine potentielle zinkbindende Domäne (Sjöberg et al., 1986).

Weitere Hinweise darauf, daß es sich bei gp sunY möglicherweise um ein DNA-bindendes Protein handelt, sind nicht bekannt.

Das sunY-Intron ist mit 1033 bp das größte bisher bekannte T4-Intron. So ist das td-Intron 1017 bp groß und das nrdB-Intron 598 bp (Belfort et al., 1985, Shub et al., 1988). Alle drei Introns beinhalten einen offenen Leserahmen. Während der offene Leserahmen nrdB il (die Bezeichnung "il" steht entsprechend der von Kutter et al., 1987 festgelegten Nomenklatur für den von einem Intron kodierten Leserahmen und bedeutet "Intron located") für ein 97 Aminosäuren großes Protein kodiert, besitzen die anderen beiden Leserahmen orf td il und orf 55.12 Kodierungskapazitäten für ein 245 bzw. 258 Aminosäuren großes Protein.

Sowohl gp td il als auch gp 55.12 sind basische Proteine. Es konnten keine Sequenzhomologien zwischen den drei von Introns kodierten T4-Proteinen festgestellt werden.Dafür läßt sich aber für orf td il als auch für orf 55.12 ein Modell für die Regulation der Genexpression formulieren, daß demjenigen für die Expression von Gen 49 gleicht. So ist bei beiden Leserahmen die Ribosomenbindestelle in dem Stamm einer ausgeprägten Sekundärstruktur enthalten. Ein später Promotor ist bei den beiden Genen jeweils so orientiert, daß bei Initiation der späten Transkription diese mRNA das Palindrom nicht mehr ausbilden kann.

Der Kodon-Gebrauch von orf 55.12 sowie von orf td il und nrdB il entspricht typischen T4-Genen. Das gilt ebenso für die Exonstrukturen von sunY, nrdB und td (Shub et al., 1988).

Ein Vergleich der sunY-Spleißsequenz mit den Nukleotidabfolgen um die Spleißstelle von td und nrdB zeigt einige homologe Sequenzabschnitte (s. Abb. 29).

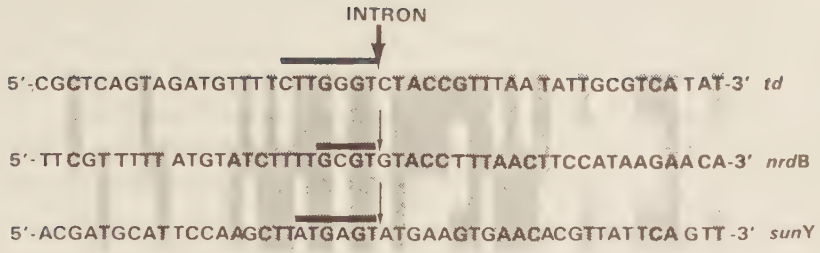


Abb. 29: Vergleich der Spleißsequenzen von td, nrdB und sunY (nach Shub et al., 1988).

Die Spleißstelle ist mit einem Pfeil gekennzeichnet. Sequenzhomologien sind schattiert abgebildet.

Ein Homologievergleich von Genprodukt sunY mit anderen Proteinen (Dayhof Datenbank, 8. Fassung) zeigt partielle Sequenzhomologien zu verschiedenen DNA- und RNA-Polymerasen: DNA-Polymerase (Bakteriophage T4), RNA-Polymerase, 160 kd Polypeptid (Saccharomyces cerevisiae), DNA-Polymerase I (E. coli) und RNA-abhängige RNA-Polymerase (Bakteriophage MS2). Hier sind in einer bestimmten 30 Aminosäuren großen Sequenzabfolge jeweils 10 identische Aminosäuren angeordnet. Da die homologen Abschnitte innerhalb von Genprodukt sunY aber an verschiedenen Bereichen lokalisiert sind, lassen sich keine Rückschlüsse auf die Funktion dieses Proteins ziehen.

Auch innerhalb der Intronsequenz von sunY finden sich konservierte kurze Sequenzen, wie sie für Introns der Gruppe I charakteristisch sind. Diese 4 ca. 10-12 Basen großen Nukleotidabfolgen werden mit P, Q, R und S bezeichnet (Waring & Davies, 1984). Es können jeweils P-Q und R-S Basenpaarungen miteinander eingehen, wobei die komplementären Sequenzen jeweils durch längere Basenabschnitte voneinander getrennt sind (s. Abb. 36 im Anhang). Zusätzlich zu diesen Basenpaarungen finden sich innerhalb der Intronsequenz noch weitere Palindrome; hierbei wird der Stamm von eng benachbart angeordneten Sequenzen ausgebildet (P3 bis P9). Die P1 und P10 Basenpaarungen finden zwischen benachbarten Intron und Exon-Sequenzen statt (Waring & Davies, 1984). P2 kommt nicht bei allen Introns vor (Burke et al., 1987).

Ein Vergleich der P, Q, R und S Sequenzen von sunY mit denen von nrdB und td sowie der Konsensussequenz für Introns der Gruppe I zeigt eine große Homologie (s. Abb. 36, Anhang).

Im Anhang in Abbildung 37 sind abschließend Sekundärstruktur-Modelle für die Introns von sunY, td und nrdB sowie eine mögliche, aus diesen drei Modellen abgeleitete Konsensussequenz für T4-Introns dargestellt.

4.2.2. Orf 49.1 ist Gen pin A

In E. coli-Zellen ist die Halbwertszeit von Polypeptidfragmenten und "abnormalen, fremden" Proteinen viel geringer als die Halbwertszeit für "normale" Proteine. Anscheinend besitzt E. coli zwei unterschiedliche proteolytische Systeme: Eine Protease-Aktivität ist für den langsamen Abbau von "normalen" Proteine verantwortlich, das andere System degradiert "abnormale" Proteine mit einer weitaus größeren Geschwindigkeit (Schoemaker, 1983; Goldberg & St. John, 1976). Wie die beiden proteolytischen Systeme zwischen normalen und abnormalen Proteinen unterscheiden ist unbekannt.

An der Degradierung von abnormalen Proteinen sind neben der lon-Protease noch weitere Enzyme beteiligt (Hwang et al., 1987; Katayama-Fujimura et al., 1987).

Werden E. coli-Zellen mit dem Bakteriophagen T4 infiziert, wird die Degradierung abnormaler Proteine inhibiert. Der Proteinabbau von normalen E. coli-Proteinen wird dagegen nicht beeinflußt (Simon et al., 1978). Diese Funktion wurde "pin" genannt (proteolysis inhibition).

Neben der lon-Protease werden wahrscheinlich auch andere E. coli-Proteasen von der pin-Aktivität inhibiert. So werden abnormale Proteine von einer lon^--Mutante in größerem Umfang abgebaut als von einem lon-Wildtyp, der die pin-Aktivität besitzt. Bei Anwesenheit von pin in lon^--Zellen ist die Degradierungsrate von abnormalen Proteinen geringer als bei lon^--Zellen ohne pin (Simon & Fay, 1982).

Hinweise darauf, daß die pin-Funktion von mehreren T4-Genprodukten ausgeübt wird, geben die Arbeiten von Skorupski et al. (1988). Hier wurde für den offenen Leserahmen 49.1 in vivo eine inhibierende Wirkung auf die lon-Protease nachgewiesen.

Im Gegensatz zu einer T4-Infektion inhibiert ein rekombinanter Vektor mit orf 49.1 zwar den Proteinabbau in E. coli-Zellen, nicht aber in lon$^-$-Mutanten. Aus diesem Grund wurde orf 49.1 als pinA bezeichnet. Vermutlich kodiert das T4-Genom neben pinA (Inhibitor der lon-Protease) noch für mindestens einen weiteren Protease-Inhibitor ("pinB"?). Über die mögliche Lage dieses Gens gibt es keine Hinweise.

Daß es sich bei orf 49.1 um pinA handelt, ist in Übereinstimmung mit früheren Ergebnissen. Wie Simon et al., 1978 zeigten, repräsentiert pin eine frühe T4-Funktion. Bei Zugabe von Chloramphenicol 1 Minute nach T4-Infektion ist keine pin-Aktivität festzustellen, während die Zugabe von Chloramphenicol 10 Minuten nach Infektion nicht mehr die Ausprägung der pin-Aktivität inhibiert. Darüberhinaus wird die pin-Funktion auch in T4 33$^-$- und 55$^-$-Mutanten synthetisiert (Genprodukt 33 und 55 werden für die Expression später T4-Gene benötigt), was ebenfalls dafür spricht, daß pin ein frühes T4-Gen ist. Bei orf 49.1 handelt es sich um ein frühes Gen, da stromaufwärts vom ATG-Initiationskodon ein starker früher Promotor lokalisiert ist. Aus der Beobachtung, daß die stabile Klonierung von starken T4-Promotoren schwierig ist, läßt sich die Instabilität eines pin-Aktivität besitzenden Klons erklären (Skorupski et al., 1988).

Auf welche Weise das Genprodukt pinA die lon-Protease von E. coli inhibiert ist unbekannt.

Der Polaritätsindex von 50 für Genprodukt pinA läßt darauf schließen, daß es sich um ein lösliches Protein handelt.

Ebenso wie T4 besitzen auch die Bakteriophagen T5 und T7 eine Protease-inhibierende Wirkung, jedoch in geringerer Ausprägung als bei T4 (Simon et al., 1978). Ein Vergleich der Aminosäuresequenz von gp pinA mit den aus der Nukleotidsequenz von T7 abgeleiteten Aminosäuresequenzen (Dunn & Studier, 1983) zeigt allerdings keine ausgedehnten Homologien.

4.2.3. Gen nrdC und orf 55.7

Die Reduktion der Ribose am C-Atom 2 verläuft in E. coli über zwei alternative Wege: sowohl reduziertes Thioredoxin als auch reduziertes Glutaredoxin sind die unmittelbaren Wasserstoffdonatoren bei der Ribosereduktion (Thelander & Reichard, 1979), welche von dem Enzym Ribonukleotid-diphosphat-Reduktase katalysiert wird.

Der Bakteriophage T4 kodiert für eine eigene Ribonukleotid-Reduktase (Genprodukte nrdA und nrdB) und für ein eigenes Thioredoxin (Genprodukt nrdC). Während die Gene nrdA und nrdB bei etwa 136 kb bis 140 kb auf dem Phagengenom positioniert sind, kartiert Gen nrdC bei 48 kb. In Anwesenheit der E. coli-Thioredoxin-Reduktase wird die singuläre Disulfidbrücke von T4-Thioredoxin unter Oxidation von $NADPH+H^+$ reduziert.

Das T4-Thioredoxin ist aus 87 Aminosäuren aufgebaut und ist somit größer als das E. coli-Glutaredoxin und kleiner als das 108 Aminosäuren große E. coli-Enzym (Wallace & Kushner, 1984; Höög et al., 1983). Die Primärstruktur dieser Proteine zeigt gewisse Sequenzhomologien (s. Abb. 30). Das aktive Zentrum von T4-Thioredoxin besteht aus den beiden Aminosäuren Cys-14 und Cys-17. Diese Sulfhydrylgruppen werden auch bei E. coli-Thioredoxin, Thioredoxin von Corynebakterium nephridii (Meng & Hogenkamp, 1981) und Glutaredoxin von E. coli (Höög et al., 1983) gefunden. Wie in Abb. 30 dargestellt ist, zeigt auch die von orf 55.7 abgeleitete Aminosäuresequenz um das potentielle aktive Zentrum Homologien zu den genannten Enzymen.

Die größte Homologie weist Genprodukt 55.7 mit T4-Thioredoxin auf, aber auch zu Thioredoxin und Glutaredoxin von E. coli ist eine partielle Homologie nachweisbar. Vielleicht handelt es sich bei Genprodukt 55.7 um ein von T4 kodiertes Glutaredoxin.

Die dreidimensionale Struktur von T4-Thioredoxin und die des E. coli-Enzyms wurde bestimmt; beide Enzyme besitzen eine ähnliche Tertiärstruktur (Holmgren et al., 1975; Eklund et al., 1984).

Das Gen nrdC konnte ohne Schwierigkeiten in den Expressionsvektor pT7-6 kloniert und zur Expression gebracht werden. Drei Stunden nach der Induktion beträgt der Anteil von T4-Thioredoxin ca. 15-20% am gesamten Zellprotein. Die von Le Master (1986) publizierte Überexpressionsrate von ca. 5% liegt etwas unter der in dieser Arbeit erreichten Syntheserate. Wahrscheinlich

läßt sich dieser Unterschied auf die Verwendung eines anderen Expressionssystems (P_L-Promotor, Vektor pKC30) zurückführen.

```
(1)  M Q T V I F G R S - - - G C P Y C V R A - K D L A E K L S - - - - - - N E R D D F Q   E. coli Glutaredoxin

(1)  M F K V Y G Y D S N I H K C V Y C D N A - K R L L T V K K Q P F E F I N I M P E K G   T4 Thioredoxin

(5)  K I E V Y G I P D E V G R C P G C Q S V T K - L L K E L N A P F T F Y K V L T N N G   T4 gp orf 55.7

(22) A I L V D F W A E - - - W C G P C K M I - A P I L D E I A D E Y Q G K L T V A K L N   E.coli Thioredoxin

(20) P V V V D F W A E - - - W C G P C K M I - A P M I E K - A T E M A G Q V K I A K V N   C. neph. Thioredoxin
```

Abb. 30: Vergleich der Aminosäuresequenzen von Genprodukt 55.7 von T4 mit T4-Thioredoxin, Thioredoxin und Glutaredoxin von E. coli und Thioredoxin von Corynebakterium nephridii im Bereich des aktiven Zentrums (mod. nach Höög et al., 1983).

Links ist die jeweilige Position der ersten gezeigten Aminosäure innerhalb des entsprechenden Proteins angegeben. Eingerahmt sind nur die Homologien zu gp 55.7.

4.2.4. Genprodukt 49 ist Endonuklease VII

Die Position und Orientierung von Gen 49 als ein spätes T4-Gen ist ungewöhnlich, denn es kartiert zwischen den frühen Genen nrdC und 55. Außerdem wird es nicht, wie für späte Gene bei T4 üblich, im Uhrzeigersinn transkribiert. Allerdings werden auch einige andere späte T4-Gene (Gen soc und Gen e) wie Gen 49 entgegen dem Uhrzeigersinn transkribiert. Für diese Gene wird ein ähnlicher Regulationsmechanismus für die Genexpression beschrieben (Mc Pheeters et al., 1986; Macdonald et al., 1984). Daß Endonuklease VII tatsächlich eine späte T4-Funktion ist, zeigen Experimente von Kemper & Garabett (1981). Sie wiesen nach, daß die Endonuklease VII-Aktivität in Gen 55^--Mutanten um das 100-fache reduziert ist. Gen 55 wird zur Erkennung von

späten T4-Promotoren durch die von Phagenproteinen modifizierte RNA-Polymerase von E. coli benötigt. Aufgrund der Genomorganisation ist zu vermuten, daß Gen 49 auch während der frühen Phase der Phagenentwicklung transkribiert wird. Wie in den Versuchen zur Expression von Gen 49 gezeigt wurde, kann aber nur die späte mRNA aufgrund des Fehlens der Sekundärstruktur Sk-1 effizient translatiert werden.

Die Reinigung von Endonuklease VII aus T4-infizierten Zellen wurde von zwei Arbeitsgruppen unabhängig voneinander durchgeführt. Während Nishimoto et al. (1979) das Molekulargewicht der aktiven Endonuklease VII durch Gelfiltration mit etwa 50000 d bestimmten, geben Kemper und Garabett (1981) das Molekulargewicht dieses Enzyms aufgrund von Sedimentationsanalysen mit 40000 d und aufgrund von Gelfiltrationsversuchen mit 42000 d an. Da das Molekulargewicht des von der Gen 49-Nukleotidsequenz abgeleiteten Proteins 18123 d beträgt, liegt die Vermutung nahe, daß es sich bei dem aktiven Enzym um ein aus zwei identischen Untereinheiten aufgebautes Protein handelt. Für diese Hypothese spricht einerseits die symmetrische Anordnung der in vitro bestimmten Schnittpositionen an kruziformer DNA (Lilley & Kemper, 1984; Kemper et al., 1985), als auch andererseits die Tatsache, daß Rohextrakte aus Genprodukt 49 überproduzierenden Zellen eine nachweisbare Endonuklease VII-Aktivität besitzen. Versuche von Kosak (Köln) zeigen, daß ein von Gen 49 überproduziertes Protein wahrscheinlich die gleichen Enzymaktivitäten aufweist wie Endonuklease VII aus T4-infizierten Zellen. Somit ist höchstwahrscheinlich nur Genprodukt 49 am Aufbau von Endonuklease VII beteiligt.

Der Bakteriophage T7 kodiert ein Enzym, welches eine sehr ähnliche Aktivität aufweist wie Endonuklease VII von T4. Das Enzym (Endonuklease I) wird von Gen 3 kodiert. Beide Enzyme besitzen eine ausgeprägte Spezifität für einzelsträngige DNA und schneiden supercoil-DNA bevorzugt an palindromischen Sequenzen (de Massy et al., 1987). Darüberhinaus findet eine Komplementation von T4-Gen 49^--Mutanten durch Gen 3 von T7 statt (de Massy et al., 1985). Beide Proteine sind in etwa gleich groß (gp 49: 157As., gp 3: 149 As.). Ein Vergleich der Aminosäuresequenzen zeigt eine partielle Homologie (30,4% identische Aminosäuren innerhalb eines 46 Aminosäuren großen Bereichs). Die Homologie ist in Abb. 31 gezeigt.

```
          30         40         50         60         70
Gp 3  : KGIKFEYEEWKVPYVIPASNHTYTPDFLLPNGIFVETKGLWESDDRK
        ::   .: :: .  ...  . .::.. . ::  ::..:.   :   :
Gp 49 : KGQGVDYLEW-LENLLTYLKSDYTQNNIHPN--FVGDKSKEFSRLGK
         80         90        100        110        120
```

Abb. 31: Partielle Homologie zwischen Genprodukt 3 von T7 und Genprodukt 49 des Bakteriophagen T4.

Identische Aminosäuren sind mit einem Doppelpunkt gekennzeichnet, konservative Substitutionen durch einen Punkt.

Weiterhin bestehen teilweise Homologien mit Endonuklease Eco RI, gp recF und der T4-kodierten Endonuklease V (gp denV). Diese Homologien beschränken sich auf die Aminosäurereste 111-136 des Genprodukts 49 (s. Abb. 32).

```
L A V G E V S R L V K K A L S N E Y P Q L S F R Y R     Endonuklease EcoRI

L T F S F Q R G W E K E T E Y A E V L E R N F E R D     E. coli rec F

D K S K E F S R L G K E E M M A E M L Q R G F E Y N     T4 Gp 49 (Endonuklease VII)

D K - L E F L R K R Q I E L I A E C L K R G F N I K     T4 Endonuklease denV
```

Abb. 32: Homologien von Genprodukt 49 mit verschiedenen DNA-bindenden Proteinen.

Die zu Endonuklease VII identischen Aminosäurereste sind eingerahmt. Die Position der jeweils ersten gezeigten Aminosäure ist: Eco RI (32), gp recF (215), Endonuklease VII (111), Endonuklease V (62).

Genprodukt recF ist ein an der Rekombination beteiligtes Protein. So ist die verbleibende Rekombinationsrate bei recBC-Mutanten auf das recF-System zurückzuführen (Clark et al., 1984). Auch für Endonuklease VII wird eine Beteiligung an der Rekombination vermutet (Warner & Snustad, 1983).

Endonuklease Eco RI erkennt und schneidet doppelsträngige DNA innerhalb der Hexanukleotidsequenz GAATTC. Das Molekulargewicht des aktiven Dimers beträgt 62130 d (Newman et al., 1981). Das Protein wurde kristallisiert und die dreidimensionale Struktur bestimmt (Rosenberg et al., 1987). Die für die Nukleotidbindung wichtigen Aminosäuren sind Glu-144, Arg-145 und Arg-200. Diese Aminosäuren sind nicht in dem hier gezeigten, zu gp 49 homologen Proteinabschnitt lokalisiert.

Endonuklease V von T4 führt Einzelstrangbrüche in doppelsträngige DNA ein, welche Pyrimidin-Dimere besitzt (Nakabeppu & Sekiguchi, 1981).

Alle hier gezeigten Enzyme besitzen also eine DNA-bindende und eine hydrolytische Aktivität. Auffallend ist bei allen vier Proteinen ein konservierter Glutaminsäurerest und ein Phenyalaninrest. Weiterhin ist bei Eco RI, gp 49 und gp recF ein Lysinrest konserviert. Ob diese Homologien in einer für die Enzymfunktion wichtigen Region lokalisiert sind bleibt unbekannt.

Wie bei Genprodukt sunY so kann auch bei Endonuklease VII von T4 eine potentielle zinkbindende Domäne postuliert werden (s. Abb. 33).

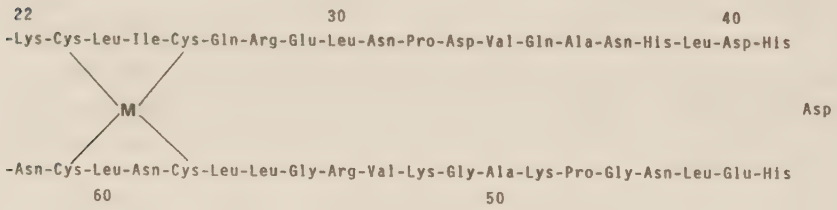


Abb. 33: Eine mögliche zinkbindende Domäne bei Endonuklease VII.

Dieser "Zink-Finger" findet sich nicht bei Endonuklease I von T7. Wie bei gp sunY baut sich auch bei gp 49 diese Struktur aus vier Cysteinresten auf. Klug & Rhodes (1987) vermuten, daß Cys-Cys····Cys-Cys Domänen möglicherweise nur an DNA binden können, während Proteinbereiche der Struktur Cys-Cys···His-His sowohl an DNA als auch an RNA binden können. Tatsächlich ist für Endonuklease VII nur eine DNase-Aktivität aber keine RNase-Aktivität beschrieben worden.

Bevor auf die Expression von Gen 49 in Abschnitt 4.4. näher eingegangen wird, sollen erst die möglichen Promotor- und Sekundärstrukturen innerhalb der 9,1 kb großen Sequenz diskutiert und ein Modell für die transkriptionelle Kontrolle dieses Genomabschnitts vorgestellt werden.

4.3. Analyse der 9,1-kb Region auf Transkriptionskontrollelemente

Auf der Sequenz sind mindestens drei mögliche frühe, 5 mittlere und 2 späte Promotoren angeordnet.

Alle drei frühen Promotoren lassen sich aufgrund ihrer Primärstruktur und des Abstandes zwischen der -35 und der -10 Box den starken frühen Promotoren zuordnen. In einem *in vitro*-Transkriptionssystem wurden von Gram et al. (1984) in dieser Region allerdings nur P_{e-2} und P_{e-3} kartiert. Folgende Gründe deuten aber auf eine Promotor-Aktivität von P_{e-1} hin:

1) Die -35 und -10 Box sowie der Abstand zwischen diesen beiden regulatorischen Sequenzen ist bei dem kartierten Promotor P_{e-3} identisch mit dem von P_{e-1}.
2) Mit einem speziellen Vektorsystem, welches zur selektiven Klonierung von starken Promotoren konzipiert ist, wurden P_{e-1} bis P_{e-3} als effektive Initiationsstrukturen für die Transkription erkannt (Liebig, 1987).
3) Alle drei frühen Promotoren konnten in M13 mp18 nur in einer Orientierung kloniert werden.

Mittlere Promotoren konnten bisher nicht in einem *in vitro*-Transkriptionssystem nachgewiesen werden, da nur mit unmodifizierter *E. coli*-RNA-Polymerase gearbeitet wurde. Die fünf potentiellen mittleren Promotorstrukturen besitzen eine relativ große Homologie zu der Konsensussequenz (s. Tab. 4). Interessanterweise überlappen die Sequenzen der beiden mittleren Promotoren P_{m-4} und P_{m-5}, welche direkt stromaufwärts von Gen 55 angeordnet sind. Der Abstand zwischen der -35 und der -10 Box von P_{m-4} mit 17 bp deutet auf einen schwachen Promotor hin, während der Abstand von 12bp bei P_{m-5} auf einen stärkeren Promotor schließen läßt.

Der späte Promotor P_{l-1} unterscheidet sich ebenso wie P_{l-2} in drei Basen von der Konsensussequenz für diese Promotorklasse. Da besonders das erste Thymidin bei den bisher bekannten Promotoren stark konserviert ist, scheint es sich bei P_{l-1}, unter dessen Kontrolle sich Gen 49 befindet, um einen schwachen Promotor zu handeln.

Auffallend ist, daß alle fünf mutmaßlichen mittleren Promotoren (ebenso wie der frühe Promotor P_{e-3}) zumindest teilweise innerhalb kodierender Sequenzen positioniert sind. Das Vorhandensein von Promotoren innerhalb von Genstrukturen ist bei T4 nicht ungewöhnlich. So ist ein starker früher Promotor innerhalb von Gen 39 angeordnet (Huang, 1986), ein weiterer befindet sich innerhalb von Gen 47 (Gram & Rüger, 1985; Liebig, 1987).

Die Sekundärstrukturen Sk-0 bis Sk-5 können nicht alle Transkriptionsterminatoren zugeordnet werden. Sk-1 und Sk-3 regulieren die Genexpression wahrscheinlich auf der Ebene der Translation (s. 4.4.). Sk-4 ist innerhalb eines Leserahmens angeordnet und fungiert somit nicht als Terminator, sondern vielleicht als eine Attenuator-ähnliche Struktur. Sk-2 entspricht den innerhalb des sunY-Introns positionierten Sekundärstrukturen P7.1 und P7.2 (s. Abb. 37). Die Haarnadelstruktur Sk-0 repräsentiert dagegen möglicherweise einen faktorabhängigen Terminator, da an ihrem Ende die für rho-unabhängige Terminatoren charakteristischen 6-8 Uridinreste fehlen. Sk-5 ist dagegen wahrscheinlich ein rho-unabhängiger Terminator. Hier befinden sich am 3´Ende sieben Uridinreste, unterbrochen durch zwei Cytidine ("UUUCUUUCU").

Nachfolgend soll ein Modell einer transkriptionellen Kontrolle der 9,1 kb großen Sequenz vorgestellt werden.

4.3.1. Modell einer möglichen Transkriptionskontrolle für die zwischen 48,266 kb und 39,166 kb positionierten Genstrukturen

Etwa 2 kb stromaufwärts von Gen nrdC ist ein starker früher T4-Promotor positioniert (Gram et al., 1984; Liebig, 1987). Vorstellbar ist, daß von hier ausgehend Gen nrdC transkribiert wird. Innerhalb des "vor" Gen nrdC gelegenen Leserahmens befindet sich zusätzlich der potentielle mittlere Promotor P_{m-1}. Durch diese Anordnung wird sichergestellt, daß in der frühen

und mittleren Phase des Infektionszyklus Gen nrdC, orf 49.1 und orf 49.2 transkribiert und exprimiert werden. Die Hypothese ist in Übereinstimmung mit Wood & Revel (1976), wonach nrdC unter der Kontrolle eines frühen oder mittleren Promotors stehen soll. NrdC, 49.3 und 49.2 sind durch ATGA-Elemente verbunden, was darüberhinaus auf eine enge translationale Kopplung schließen läßt. Der Abbruch der mRNA-Synthese für diesen Genbereich erfolgt an dem rho-abhängigen Terminator Sk-0. Hierbei sind die ersten beiden Basen des nachfolgenden frühen Promotors P_{e-1} bereits im Stamm der Haarnadelstruktur lokalisiert.

Von P_{e-1} aus wird eine etwa 5,6 kb große mRNA synthetisiert. Die frühe mRNA kodiert für die Gene pinA, 49, sunY und orf 55.10 bis 55.7. Diese Anordnung bedeutet aber nicht, daß alle Gene auch in der frühen Phase des T4-Infektionszyklus exprimiert werden. Sowohl bei Gen 49 als auch dem innerhalb des Introns befindlichen Leserahmen 55.12 wird durch das Vorhandensein der jeweiligen Ribosomenbindestelle im Stamm einer ausgedehnten Sekundärstruktur (Sk-1 bzw. Sk-3) die effiziente Translation zu diesem Zeitpunkt verhindert.

Der innerhalb von Gen 49 vorhandene mittlere Promotor P_{m-2} soll möglicherweise gewährleisten, daß sunY auch während der mittleren Phase der Phagenentwicklung transkribiert und exprimiert wird. Auch orf 55.8 und 55.7 können zu diesem Zeitpunkt ausgehend von P_{m-3} transkribiert werden.

Während der späten Infektionsphase wird die mRNA von P_{l-1} und P_{l-2} synthetisiert, was zur Expression von Endonuklease VII und gp 55.12 führt.

Der Abbruch der mRNA-Synthese findet an Sk-5 statt. Die Leserahmen 55.6 bis 55 werden von den frühen Promotoren P_{e-2} und P_{e-3} aus transkribiert. Der direkt stromaufwärts von Gen 55 gelegene Tandempromotor P_{m-4}/P_{m-5} gewährleistet zusätzlich eine Transkription von Gen 55 auch während der mittleren Phase der T4-Infektion.

Die Positionen der offenen Leserahmen, der Promotoren und der Sekundärstrukturen sind der Abb. 34 (s. Anhang) zu entnehmen.

4.4. Expression von Endonuklease VII

Während das Gen nrdC problemlos in das pT7-System kloniert und zur Expression gebracht wurde, gelang eine Überexpression von Gen 49 erst nach Deletion von Sk-1 mit anschließender Klonierung dieses Fragments in einen modifizierten pT7-Vektor (pJT7-6TpAT8).

Zwar kann ein DNA-Fragment mit Gen 49 und orf 55.13 in pT7-Plasmide kloniert werden, aber es kommt nach einer Induktion nur zur Überproduktion von gp 55.13. Daß es sich bei dem synthetisierten 20 kd großen Protein nicht um Genprodukt 49 sondern um gp 55.13 handelt, belegen die in Kapitel 3.4.3. beschriebenen Subklonierungen. Warum wird dieses Protein in einer Konzentration von mehr als 30% des Zellproteins synthetisiert? Der Kodongebrauch entspricht weitgehend dem eines typischen T4-Gens und die Ribosomenbindestelle ist im Gegensatz zu Gen nrdC oder Gen 49 nur schwach ausgeprägt. Orf 55.13 repräsentiert das erste Exon des Mosaikgens sunY. Das in pT7-5 bzw. pT7-6 klonierte DNA-Fragment beherbergt neben Gen 49 und orf 55.13 noch etwa 190 bp des Introns einschließlich der Sequenzen, welche die Intron-Sekundärstrukturen P1, P3, P4, P5, P6, P7, P7.1, P7.2 und P8 ausbilden können (s. Abb. 37). Für die ausgehend von Ø10 synthetisierte mRNA lassen sich somit für das 3´Ende eine Vielzahl partiell doppelsträngiger Bereiche vorhersagen. Vielleicht verhindern diese Sekundärstrukturen einen möglichen Abbau der mRNA vom 3´Ende aus. Andererseits wird aber auch das aminoterminale 12 kd-Proteinfragment des Genprodukts 55.13 von dem Klon pJT7-5/49A in großer Konzentration überproduziert, ohne daß hier die oben genannten Haarnadelstrukturen beteiligt sein können.

Gen 49 und orf 55.13 sind durch ein ATGA-Element verbunden. Diese Kopplung ist aber anscheinend nicht für die effiziente Expression von orf 55.13 notwendig, da gp 55.13 in etwa in gleich großer Konzentration von pJT7-6/orf 55.13 und von pJT7-5/49 synthetisiert wird. Dem Klon pJT7-6/orf 55.13 fehlen die ersten 190 bp von Gen 49; somit kann keine Translation von Gen 49 initiiert werden und die Ribosomen müssen an der schwachen Ribosomenbindestelle von orf 55.13 die Proteinbiosynthese neu initiieren. Allerdings kann unter der Voraussetzung, daß an einem innerhalb von Gen 49 lokalisierten offenen Leserahmen <u>in vivo</u> translatiert wird, die Theorie der translationalen Kopplung aufrecht erhalten werden, denn die potentielle Ribosomenbindestelle von orf 49** (AAGG, 4 bp Abstand zum ATG-Kodon)ist bei dem Klon

pJT7-6/orf 55.13 noch vorhanden.

Die Deletion der Sekundärstruktur Sk-1 erhöht die Expression von Gen 49 um etwa das Zehnfache (s. 3.4.8.), auch wenn diese Plasmide nicht in E. coli K38/pGP1-2 transformiert werden können. Die Steigerung der Proteinproduktion durch Deletion von Sk-1 untermauert die in Abschnitt 3.4.3. aufgestellte Hypothese, wonach in der frühen und späten Phase der T4-Infektion Gen 49 mRNA synthetisiert wird, aber nur die späte mRNA aufgrund des Verlustes von Sk-1 effizient translatiert werden kann.

Dieser Regulationsmechanismus scheint auch bei anderen T4-Genen realisiert zu sein, so z.B. bei Gen soc (Macdonald et al., 1984), bei Gen e (Mc Pheeters et al., 1986) und bei den Intron-kodierten Genen td il (Belfort et al., 1985) und orf 55.12 (Shub et al., 1988). Im Falle des für T4-Lysozym kodierenden Gens e wurde nachgewiesen, daß eine Translation an der frühen mRNA aufgrund der direkt stromaufwärts positionierten Sekundärstruktur inhibiert ist (Mc Pheeters et al., 1986). Hierbei ist, wie bei Gen 49, in dem Stamm der Haarnadelstruktur die Ribosomenbindestelle angeordnet. Die späte mRNA wird bei Gen e etwa 100 mal stärker translatiert als die frühe mRNA (Kasai & Bautz, 1969). Werden allerdings in den Stamm des Palindroms Mutationen dergestalt eingeführt, daß die Stabilität der Sekundärstruktur erniedrigt wird, so erhöht sich in der frühen Phase der T4-Infektion die Expression des Lysozym-Gens (Knight et al., 1987). Die für Gen e beobachteten Ergebnisse sind ohne Einschränkung auf Gen 49 übertragbar, wie die in dieser Arbeit dargestellten Experimente zur Überproduktion von Endonuklease VII zeigen.

Eine Genregulation der oben beschriebenen Form ist aber nicht allein auf T4 beschränkt. Die Expression der Tn10-Transposase erfolgt nach einem ähnlichen Schema. Wird ausgehend von Wirtspromotoren in Richtung des inserierten Transposons transkribiert, bildet die mRNA eine stabile Sekundärstruktur aus, bei der das ATG-Initiationskodon im Stamm "kaschiert" ist. Hierdurch wird die Expression der Transposase vollständig inhibiert. Startet die Transkription dagegen von einem bestimmten, stromaufwärts des Palindroms angeordneten Tn10-Promotor aus, so fehlt dieser mRNA die Sekundärstruktur, und das Transposase-Gen wird translatiert (Davis et al., 1985).

Die Endonuklease VII-Aktivitätstests belegen, daß die Blockierung der Gen 49-Expression durch Sk-1 im pT7-System nicht vollständig ist. Somit ist nicht auszuschließen, daß auch bei T4-infizierten Zellen in der frühen Phase der Phagenentwicklung eine geringe Menge an Endonuklease VII synthetisiert wird.

Da für gp 49 eine Beteiligung an der DNA-Replikation vermutet wird (Mosig, 1983), ist vorstellbar, daß die wenigen Nuklease-Moleküle zu diesem frühen Zeitpunkt für die T4-DNA-Synthese benötigt werden.

Die Deletion der Sekundärstruktur Sk-1 führt nicht zu einer auf Polyacrylamidgelen nachzuweisenden Überproduktion von Endonuklease VII. Nur das von pJT7-4/41-3-15M synthetisierte Protein läßt sich problemlos überproduzieren. Der Klon kodiert für ein inaktives Genprodukt 49, wie sowohl die Enzymaktivitätstests als auch die nach Gelelektrophorese nicht zu beobachtende DNA-Degradierung belegen. Das gilt ebenso für den Klon pJT7-4/41-19, der für ein etwa 8 kd großes aminoterminales Proteinfragment von gp 49 kodiert. Beide Proteine besitzen keine Enzymaktivität und lassen sich ohne Schwierigkeiten im pT7-System überproduzieren. Vermutlch liegen deshalb die Probleme, Endonuklease VII in großen Konzentrationen zu synthetisieren, in der Aktivität dieses Enzyms begründet. Dafür sprechen auch die Ergebnisse von de Massy et al. (1987). Hier konnte die klonierte Endonuklease I des Phagen T7 nicht in den Mengen überproduziert werden, um das Protein auf SDS-Gelen nachzuweisen. Wahrscheinlich sind die DNA-Schädigungen, welche beide Nukleasen *in vivo* in den Zellen verursachen, für die geringe Expression verantwortlich.

Um diese Schwierigkeiten zu umgehen, wurde das pT7-Expressionssystem so modifiziert, daß eine transkriptionsterminierende Struktur stromaufwärts des Ø10-Promotors inseriert wurde. Durch den Terminator sollte die Transkription in Richtung einklonierter Gene ohne vorherige Induktion verhindert werden.

Die verwendete poly(dA)·poly(dT)-Sequenz scheint tatsächlich die Genexpression ohne vorherige Induktion zu blockieren. Aktivitätsmessungen des Klons *E. coli* K38(pGP1-2)/pJT7-6TpAT8/41-8-4 ergaben einen Unterschied um mehr als den Faktor 100 zwischen dem uninduzierten und dem induzierten Zustand (H. Kosak & B. Kemper, pers. Mitteilung). Die noch vorhandene geringe Menge an Endonuklease VII auch in diesen Zellen ist wahrscheinlich auf eine Transkription durch die wenigen T7-RNA-Polymerase-Moleküle zurückzuführen, die aufgrund einer Transkription ausgehend des P_L-Promotors des Plasmids pGP1-2 synthetisiert werden. Denn während Plasmide mit Gen 49 ohne Sk-1 und ohne poly(dA)·poly(dT)-Terminator nicht stabil in Zellen mit pGP1-2 transformiert werden können, ist Gen 49 ohne Sk-1 problemlos in pJT7-6TpAT8 zu klonieren und das rekombinante Plasmid in *E. coli* K38/pGP1-2 zu transformieren. Für diese Hypothese spricht weiterhin, daß bei Zellen mit dem rekombinanten Plasmid pJT7-6TpAT8/41-8-4 ohne pGP1-2 kein DNA-Abbau nachzuweisen ist. Erst nach

Transformation in E. coli K38/pGP1-2 sind DNA-Abbauprodukte zu beobachten.

Werden rekombinante pT7-Plasmide mit Gen 49 ohne Sk-1 in E. coli K38/pGp1-2 transformiert, addiert sich wahrscheinlich die Menge an Endonuklease VII, welche einerseits durch Transkription von pT7-Startstellen, als auch andererseits durch Transkription von den wenigen T7-RNA-Polymerase-Molekülen resultiert, zu einer für die Zellen toxischen Konzentration. So wurden nach Transformation des Ligationsansatzes pJT7-4/41-3 in E. coli K38/pGP1-2 etwa ein Tausendstel der Klone isoliert, welche nach Transformation des gleichen Ansatzes in E. coli-Zellen ohne pGP1-2 erhalten wurden (s. 3.4.5.). Von den vier Klonen, welche ein Insert enthalten, synthetisiert nur einer ein 18 kd großes Protein, bei dem es sich um inaktive Endonuklease VII handelt. Anscheinend besitzen in E. coli K38/pGP1-2 diejenigen Zellen einen Selektionsvorteil, bei denen eine Mutation zu einer Reduzierung oder Blockierung der Expression aktiver Endonuklease VII führt. Solche Mutationen können zum Beispiel innerhalb von Gen 49 kartieren, aber auch im Gen der T7-RNA-Polymerase oder in der Sequenz des T7-Promotors Ø10 positioniert sein.

Eine Sequenzierung der T4-Mutante am E727 (49^-) von Barth et al. ergab, daß die amber-Mutation im Triplett 151 kartiert (CAG ——> TAG), was zu einem nahezu vollständigen Verlust der katalytischen Aktivität führt (Mosig, pers. Mitteilung; Kemper & Janz, 1976). Somit sind die "letzten" sechs Aminosäuren für die Enzymaktivität notwendig. Vielleicht kartiert auch die Mutation des Klons pJT7-4/41-3-15M in diesem Bereich. Das würde erklären, warum das überproduzierte Genprodukt zwar inaktiv ist, aber keine erhebliche Reduktion des Molekulargewichts aufweist.

Die hier diskutierten Ergebnisse implizieren eine weitere Anwendungsmöglichkeit des pT7-Systems, und zwar die Einführung von Mutationen in Gene, welche für E. coli toxisch sind. Hierzu wird das entsprechende rekombinante pT7-Plasmid in Zellen mit pGP1-2 transformiert und die wenigen überlebenden Klone auf eine veränderte Proteinexpression und Enzymaktivität getestet.

4.5. Vergleich der Genregionen um pinA und 49 von T2, T4 und T6

Die Gene pinA und 49 von T2 und T6 wurden über ihre Sequenzhomologie zu den entsprechenden T4-Genen isoliert und sequenziert.

Für die Regulation der Genexpression von Gen 49 bei T2 und T6 kann ein analoges Modell wie für T4 aufgestellt werden. Sowohl bei T2 als auch bei T6 sind alle Kontrollelemente für die Genexpression vorhanden, wie sie auch bei T4 in dieser Region vorkommen. Zwar wurde der frühe Promotor stromaufwärts von pinA der Bakteriophagen T2 und T6 nicht sequenziert. Die Beobachtung, daß kein Klon isoliert werden konnte, durch dessen Sequenzierung der Bereich bestimmt wurde, wo bei T4 der Promotor P_{e-1} positioniert ist, deutet aber auf die Anwesenheit eines starken Promotors hin. Bei T6 ist außerdem die -10 Box des zu T4 P_{e-1} homologen Promotors bestimmt worden. Auch bei Gen 49 von T2 und T6 ist die Ribosomenbindestelle innerhalb einer stabilen Haarnadelstruktur verborgen; stromaufwärts des Palindroms ist ebenfalls ein potentieller später Promotor angeordnet. Außerdem ist, wie bei T4, in dem für den carboxylterminalen Bereich von gp 49 kodierenden Sequenzabschnitt ein mittlerer Promotor lokalisiert.

Die Gene pinA und 49 sind bei allen drei Phagentypen sehr homolog zueinander. Beide Gene sind bei T2, T4 und T6 gleich lang, die Ribosomenbindestellen sind identisch. Ein Vergleich zeigt, daß es nur geringe Unterschiede zwischen den einzelnen Sequenzen gibt. Sowohl bei Gen 49 als auch pinA sind jeweils 13 Tripletts unterschiedlich. Auffallend ist besonders die Position 146 - 149 bei pinA. Während bei T2 und T4 die Aminosäuren Arg-Met-Ile-Gln vorkommen, besitzt das T6-Enzym die Abfolge Gln-Lys-Ile-Arg. Hier ist der Argininrest mit einem Glutaminrest vertauscht. Der Austausch von Methionon durch Lysin führt zu einer zusätzlichen positiven Ladung des T6-Enzyms, hat aber keinen wesentlichen Einfluß auf das Hydropathieprofil des Enzyms.

Die Endonuklease VII von T2 und T6 besitzt ebenso wie bei T4 eine mögliche metallatombindende Proteindomäne (s. 4.2.4.).

Das Gen pinA von T4 weist eine größere Homologie zu dem entsprechenden Gen von T2 und T6 auf als die Leserahmen von T2 und T6 untereinander. Bei Gen 49 ist der Unterschied zwischen T4 und T6 größer als der Unterschied zwischen T2 und T4. Die Sequenz der α-Glukosyltransferasegene von T2, T4 und T6 läßt dagegen auf einen größeren Verwandtschaftsgrad von T4 und T6 schließen (Gram & Rüger, 1986).

4.6. Vergleich der T2- und T6-Nukleotidsequenzen mit der Spleißstelle des T4-Mosaikgens sunY

Bei den Bakteriophagen T2 und T6 fehlt die intervenierende Sequenz (s. Abb. 26). Da die Nukleotidabfolgen von T2 und T6 in dem gezeigten Bereich (33 bp) identisch sind, werden sie im weiteren Verlauf als "T2/T6"-DNA bezeichnet.

Die T2/T6-DNA besitzt eine sehr große Homologie in der Spleißregion zu der von der reifen mRNA des Mosaikgens sunY abgeleiteten DNA-Sequenz. Bei sieben Nukleotidaustauschen in diesem Abschnitt ist bei T2/T6 nur eine Aminosäure unterschiedlich: ein Phenylalaninrest ist bei T2/T6 durch einen Histidinrest ersetzt.

Während der in unserer Arbeitsgruppe untersuchte T2-Phage kein Intron in der sunY-Region besitzt, trägt ein in der Arbeitsgruppe von M. Belfort untersuchter Phagenstamm vom Typ 2 das zu T4 äquivalente Intron (M. Belfort, pers. Mitteilung). Von Belfort et al. wurde die RNA-Sequenz stromaufwärts der Spleißstelle dieses T2-Phagen erstellt; sie entspricht in dem in Abb. 26 gezeigten Bereich der T4-Sequenz. Die RNA-Sequenz von T6 wurde ebenfalls erstellt; sie entspricht der in dieser Arbeit bestimmten DNA-Sequenz von T2/T6.

Die T2- und T4-Phagen, welche das sunY-Intron besitzen, werden im folgenden als "T2 In^+" und "T4 In^+" bezeichnet, die T2- und T6-Phagentypen ohne Intron entsprechend als "T2 In^-" und "T6 In^-".

Ein Vergleich der Sequenzen im Spleißbereich ergibt, daß T2 In^- homolog zu T6 In^- ist. Ebenso ist die Nukleotidabfolge von T2 In^+ im Breich der Spleißstelle identisch zu der von T4 In^+. Wahrscheinlich erfordert der Besitz eines Introns eine bestimmte konservierte Primärstruktur im angrenzenden Exonbereich.

Die hier diskutierten Ergebnisse, wonach die nahe verwandten geradzahligen T-Phagen sich durch den Besitz bzw. den Nichtbesitz eines Introns voneinander unterscheiden, sind in Übereinstimmung mit den von Pedersen-Lane & Belfort (1987) publizierten Resultaten. Während alle drei Phagentypen das td-Intron besitzen, ist nur das T4-nrdB-Gen durch eine intervenierende Sequenz unterbrochen.

Wie kann das unterschiedliche Vorkommen von Introns bei den geradzahligen T-Phagen interpretiert werden? Pedersen-Lane und Belfort (1987) vermuten, daß diese Ergebnisse auf ein Intron-Mobilität hindeuten. Die Hypothese,

wird folgendermaßen begründet: Innerhalb der Intron-Gruppe I bei Eukaryonten sind mehrere Beispiele beschrieben, wo sich nahe verwandte Spezies durch das Vorhandensein von Introns unterscheiden, so z.B. bei der großen rRNA von Tetrahymena (Nielsen & Engberg, 1985). Sogar verschiedene Stämme derselben Spezies können sich durch den Besitz eines Introns unterscheiden: die mitochondriale große rRNA von Saccharomyces cerevisiae ist nicht bei allen Subspezies gespleißt (Jacquier & Dujon, 1985). Weiterhin lassen Versuche von Lang (1984) und Michel & Cummings (1985), bei denen ein möglicher Transfer von Introns der Gruppe I zwischen den mitochondrialen Genomen von Pilzen und der im Zellkern von Tetrahymena lokalisierten DNA festgestellt wurde, auf eine potentielle Intron-Mobilität schließen. In diesem Zusammenhang ist interessant, daß der innerhalb des td-Introns lokalisierte Leserahmen Homologien zu den Intron-kodierten Leserahmen dreier mitochondrialer Gene von Pilzen aufweist (Michel & Dujon, 1986).

Abschließend soll darauf hingewiesen werden, daß das unterschiedliche Vorkommen von Introns in homologen Genen bei T2, T4 und T6 zeigt, daß intervenierende Sequenzen als Regulationselemente der Genexpression bei nrdB und sunY nicht essentiell sind.

An der weiteren Charakterisierung von T2 In^- und T2 In^+ wird zur Zeit in Zusammenarbeit mit dem Labor von M. Belfort gearbeitet.

5. ZUSAMMENFASSUNG

1) Die Sequenz einer 9,1 kb großen DNA-Region des Bakteriophagen T4 aus den Positionen 48,266 kb bis 39,166 kb auf der physikalischen Karte wurde bestimmt und die Gene 55, 49 und nrdC identifiziert.

2) Nach Klonierung des Gen nrdC in einen Expressionsvektor konnte das 10 kd große Protein in einer Konzentration von etwa 15 - 20% des Zellproteins überproduziert werden.

3) Eine stromaufwärts von Gen 49 lokalisierte Sekundärstruktur inhibiert die Überexpression des Gens. Die Haarnadelstruktur wurde deletiert und ein Set an DNA-Fragmenten konstruiert, welche Gen 49 mit verschiedenen Ribosomenbindestellen tragen.

4) Für Genprodukt 49 konnte *in vivo* eine endonukleolytische Aktivität nachgewiesen werden. Da das benutzte Expressionssystem auch im uninduzierten Zustand zu "leaky" (durchlässig) war, ließ sich gp 49 aufgrund dieser Aktivität nur in geringer Menge synthetisieren.

5) Ein im Rahmen dieser Arbeit konstruierter Expressionsvektor, welcher weniger durchlässig ist, erlaubte die Überproduktion von gp 49 in größeren Konzentrationen. Weiterführende Experimente zeigten, daß das klonierte gp 49 wahrscheinlich die gleichen Enzymaktivitäten aufweist wie Endonuklease VII aus T4-infizierten Zellen.

6) In Zusammenarbeit mit dem Labor von Simon konnte nachgewiesen werden, daß orf 49.1 für einen Inhibitor der lon-Protease von *E. coli* kodiert (pinA).

7) Gen 49 und pinA wurden bei den zu T4 nahe verwandten Phagen T2 und T6 sequenziert.

8) Zusammen mit den Arbeitsgruppen von Shub und Belfort wurde auf der 9,1 kb großen Sequenz von T4 ein 1033 bp großes Intron gefunden und die genaue Spleißstelle bestimmt.

9) Weiterführende Versuche, bei denen bei T2 und T6 die Abwesenheit des entsprechenden Introns festgestellt wurde, sichern die unter Punkt 8 beschriebenen Ergebnisse ab.

6. LITERATURVERZEICHNIS

Abremski, K. & S. Gottesmann (1982):
Purification of the bacteriophage λ xis gene product required for λ excisive recombination.
J. Biol. Chem. 257: 9658-9662.

Amundsen, S.K., A.F. Taylor, A.M. Chaudhury, G.R. Smith (1986):
recD: the gene for an essential third subunit of Exonuclease V.
Proc. Natl. Acad. Sci. USA 83: 5558-5562.

Anraku, N. & J. Tomizawa (1965):
Molecular mechanisms of genetic recombination in bacteriophages. V. Two kinds of joining of parental DNA molecules.
J. Mol. Biol. 12: 805-815

Belfort, M., Pedersen-Lane, J., West, D., Ehrenmann, K., Maley, G., Chu, F., Maley, F. (1985):
Processing of the intron containing thymidylate synthase (td) gene of phage T4 is at the RNA level.
Cell 41: 375-382

Benton, W.D. & R.W. Davies (1977):
Screening λgt Recombinant Clones by Hybridization to Single Plaques in situ.
Science 196: 180-182

Berg, J. M. (1986):
Potential Metal-Binding Domains in Nucleic Acid Binding Proteins.
Science 232: 485-487

Biggin, M.D., Gibson, T.J., Hong, G.F. (1983):
Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination.
Proc. Natl. Acad. Sci. USA 80: 3963-3965

Birnboim, H.C. & Doly, J. (1979):
A rapid alkaline extraction procedure for screening recombinant plasmid DNA.
Nucl. Acids Res. 7: 1513-1523

Broker, T. (1973):
An electron microscopic analysis of pathways for bacteriophage T4 DNA recombination.
J. Mol. Biol. 81: 1-16

Burke, J.M., Belfort, M., T.R. Cech, R.W. Davies, R.J. Schweyen, D.A. Shub, J.W. Szostak, H.F. Tabak (1987):
Structural conventions for group I Introns.
Nucl. Acids Res. 15: 7217-7221

Capaldi, R.A. & G. Vanderkooi (1972):
The low polarity of many membrane proteins.
Proc. Natl. Acad. Sci. USA 69: 930-932

Cech, T.R. (1986):
The generality of self-splicing RNA: relationship to nuclear mRNA splicing.
Cell 44: 207-210

Clark, A.J. & A. Margulies (1965):
Isolation and characterization of recombination-deficient mutants of *Escherichia coli* K12.
Proc. Natl. Acad. Sci. USA 53: 451-459

Clark, A.J., S.J. Sandler, K.D. Willis, C.C. Chu, M.A. Blanar, S.T. Lovett (1984):
Genes of the recE and recF pathways of conjugational recombination in Escherichia coli.
Cold Spring Harbor Symp. Quant. Biol. 49: 453-462

Cohen, S.N., Chang, A.C.Y., Hsu, L. (1972):
Nonchromosomal antibiotic resistance in bacteria: Genetic transformation of Escherichia coli cells.
Proc. Natl. Acad. Sci. USA 69: 2110-2114

Cox, M.M. & I.R. Lehmann (1981):
recA protein of E. coli promotes branch migration, a kinetically distinct phase of DNA strand exchange.
Proc. Natl. Acad. Sci. USA 78: 3433-3437

Craig, N.L. & N. Kleckner (1986):
Transposition and site-specific recombination.
In: Escherichia coli and Salmonella typhimurium, (eds. J.L. Ingraham, K. Brooks Low, B. Magasanik, M. Schaechter, H. Edwin Umbarger), American Socierty for Microbiology, Washington D.C., 1054-1070

Dagert, M. & Ehrlich, S.D. (1979):
Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells.
Gene 6, 23-28

Davis, M.A., Simons, R.W., Kleckner, N. (1985):
Tn10 protects itself at two levels from fortuitous activation by external proteins.
Cell 43: 379-387

Deininger, P.L. (1983):
Random subcloning of sonicated DNA: Application to shotgun DNA sequence analysis.
Anal. Biochem. 129: 216-223

De Massy, B., Studier, F.W., Dorgai, L., Appelbaum, E., Weisberg, R.A. (1984):
Enzymes and Sites of Genetic Recombination: Studies with Gene-3 Endonuclease of Phage T7 and with site-affinity Mutants of Phage λ.
Cold Spring Harbor Symp. Quant. Biol. 49: 715-726

De Massy, B. & R.A. Weisberg (1987):
Gene-3 Endonuclease of Bacteriophage T7 resolves conformationally branched structures in double-stranded DNA.
J. Mol. Biol. 193: 359-376

Dressler, D. & H. Potter (1982):
Molecular mechanisms in genetic recombination.
Annu. Rev. Biochem. 51: 727-761

Dunn, J:J: & Studier, F.W. (1983):
Complete Nucleotide Sequence of bacteriophage T7 DNA and the Locations of T7 Genetic Elements.
J. Mol. Biol. 166: 477-535

Ehrenmann, K., J. Pedersen-Lane, D. West, R. Herman, F. Maley, M. Belfort, (1986):
Processing of phage T4 td-encoded RNA is analogous to the eukaryotic groupI splicing pathway.
Proc. Natl. Acad. Sci. USA 83: 5875-5879

Eklund, H., C. Cambillau, B.-M. Sjöberg, A. Holmgren, H. Jörnvall, J.-O. Höög, C.I. Bränden (1984):
Conformational and functional similarities between glutaredoxin and thioredoxins.
EMBO J. 3: 1443-1449

Frankel, F.R., M.L. Batcheler, C.F. Clark (1971):
The role of gene 49 in DNA replication and head morphogenesis in bacteriophage T4.
J. Mol. Biol. 62: 439-463

Fukada, K. & Abelson, J. (1980):
DNA sequence of a T4 transfer RNA gene cluster.
J. Mol. Biol. 139: 377-391

Giedroc, D.P., K.M. Keating, K.R. Williams, W.H. Konigsberg, J.E. Coleman (1986).
Gene-32 protein, the single-stranded DNA binding protein from bacteriophage T4, is a zink metalloprotein.
Proc. Natl. Acad. Sci. USA 83: 8452-8456

Goldberg, A.L. & A.C. St. John (1976):
Intracellular protein degradation in mammalian and bacterial cells, part2.
Annu. Rev. Biochem. 45: 747-803

Gott, J. Shub, D.A., M. Belfort (1986):
Multiple self splicing Introns in bacteriophage T4: Evidence from autocatalytic GTP labeling of RNA in vitro.
Cell 47: 81-87

Gottesman, M.E. & R.A. Weisberg (1970):
Prophage insertion and excision.
In: The bacteriophage Lambda (ed. A.D. Hershey), pp 113-138. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Gram, H. & Rüger, W. (1985):
Genes 55, αgt, 47 and 46 of bacteriophage T4: the genomic organization as deduced by sequence analysis.
EMBO J. 4: 257-264

Gram, H., Liebig, H.-D., Hack, A., Niggemann, E., Rüger, W. (1984):
A physical map of bacteriophage T4 including the positions of strong promotors and terminators recognized in vitro.
Mol. Gen. Genet. 194: 232-240

Gram, H. & Rüger, W. (1986):
The α-glucosyltransferases of bacteriophages T2, T4 and T6: A comparison of their primary structures.
Mol. Gen. Genet. 202: 467-470

Greene, G.L., P. Gina, M. Waterfield, A. Baker, Y. Hort, J. Shine (1986):
Sequence and Expression of Human Estrogen Receptor Complementary DNA.
Science 231: 1150-1154

Grindley, N.D.F. & R.R. Reed (1985):
Transpositional recombination in prokaryotes.
Annu. Rev. Biochem. 54: 863-896

Grosjean, H. & Fiers, W. (1982):
Preferential codon-usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes.
Gene 18: 199-209

Guerrini, F. (1969):
On the asymmetry of λ integration sites.
J. Mol. Biol. 46: 523-542

Hahn, S. (1986):
Klonierung und molekulargenetische Charakterisierung von Gen 33 und benachbarten Genen des Bakteriophagen T4.
Dissertation an der Abteilung für Biologie der Ruhr-Universität Bochum

Holmgren, A., Söderberg, B.-O., Eklund, H., Bränden, C.-I. (1975):
Three-dimensional structure of Escherichia coli thioredoxin S_2 to 2,8 Å resolution.
Proc. Natl. Acad. Sci. USA 72: 2305-2309

Höög, J.-O., Jörnvall, H., Holmgren, A., Carlquist, M., Persson, M. (1983):
The primary structure of Escherichia coli glutaredoxin.
Eur. J. Biochem. 136: 223-232

Horii, T., T. Ogawa, H. Ogawa (1980):
Organization of the recA gene of Escherichia coli.
Proc. Natl. Acad. Sci. USA 77: 313-317

Hsu, P.L. & Landy, A. (1984):
Resolution of synthetic att-site Holliday-structures by the integrase protein of bacteriophage lambda.
Nature (London) 311: 721-726

Huang, W.M. (1986):
Nucleotide sequence of a type II DNA topoisomerase. Bacteriophage T4 gene 39.
Nucleic Acids Res. 14: 7751-7765

Hwang, B.J., W.J. Park, C.H. Chung, A.L. Goldberg (1987):
Escherichia coli contains a soluble ATP-dependent protease (Ti) distinct from protease La.
Proc. Natl. Acad. Sci. USA 84: 5550-5554

Jacquier, A. & B. Dujon (1985):
An Intron-encoded protein is active in a gene conversion process that spreads an Intron into a mitochondrial gene.
Cell 41: 383-394

Jensch, F. & B. Kemper (1986):
Endonuclease VII resolves Y-junctions in branched DNA in vitro.
EMBO. J. 5, 181-189

Kasai, T. & E.K.F. Bautz (1969):
Regulation of gene-specific RNA synthesis in bacteriophage T4.
J. Mol. Biol. 41: 401-417

Kassavetis, G.A. & Geiduschek, E.P. (1984):
Defining a bacteriophage T4 late promoter: bacteriophage T4 gene 55 protein suffices for directing late promoter recognition.
Proc. Natl. Acad. Sci. USA 81: 5101-5105

Katayama-Fujimura, Y., S. Gottesman, M.R. Maurizi (1987):
A multiple-component, ATP-dependent protease from Escherichia coli.
J. Biol. Chem. 262: 4477-4485

Kemper, B. & D.T. Brown (1976):
Function of gene 49 of bacteriophage T4: Analysis of intracellular development and the structure of very-fast-sedimenting DNA.
J. Virol. 18: 1000-1015

Kemper, B. & Janz, E. (1976):
Function of gene 49 of bacteriophage T4: Isolation and biochemical characterization of very-fast-sedimenting DNA.
J. Virol. 18: 992-999

Kemper, B. & Garabett, M. (1981):
Studies on T4-head maturation. 1. Purifcation and characterization of gene-49-controlled endonuclease.
Eur. J. Biochem. 115: 123-131

Kemper, B., F. Jensch, M.v. Depka-Prondzynski, H.-J. Fritz, U. Borgmeyer, K. Mizuuchi (1984):
Resolution of Holliday-Structures by Endonuclease VII as observed in interactions with cruciform DNA.
Cold Spring Harbor Symp. Quant. Biol. 49: 815-825

Kikuchi, Y. & H.A. Nash (1979):
Nicking-closing activity associated with bacteriophage λ int gene product.
Proc. Natl. Acad. Sci. USA 76: 3760-3764

Kim, J.S. & Davidson, N. (1974):
Electron microscope heteroduplex studies of sequence relations of T2, T4 and T6 Bacteriophage DNA´s.
Virology 57: 93-111

Klug, A. & Rhodes, D. (1987):
"Zinc fingers": a novel protein motif for nucleic acid recognition.
TIBS 12: 464-469

Knight, J.A., L.W. Hardy, D. Rennell, D. Herrick, A.R. Poteete (1987):
Mutations in an upstream regulatory sequence that increase expression of the bacteriophage T4 lysozyme gene.
J. Bacteriol. 169: 4630-4636

Kotewicz, M., E. Grzesiuk, W. Courschesne, R. Fischer, H. Echols (1980):
Purification and characterization of the integration protein specified by bacteriophage λ.
J. Biol. Chem. 255: 2433-2439

Krebbers, E.T., Larrinua, I.M., McIntosh, L., Bogorad, L. (1982):
The maize chloroplast genes for the ß- and α-subunit of the photosynthetic coupling factor CF-1 are fused.
Nucl. Acids res. 10: 4985-5002

Kutter, E. & Rüger, W. (1983):
Map of the T4 genome and its transcriptional control sites.
In: the bacteriophage T4, (eds. Mathews, C.K., Kutter, E.M., Mosig, G., Berget, P.B.), pp 277-290, American Society for Microbiology, Washington, D.C.

Kutter, E., Guttman, B., Rüger, W., Tomaschewski, J., Mosig, G. (1987):
Bacteriophage T4.
In: Genetic maps 1987. Acompilation of linkage and restriction maps of genetically studied organisms. Vol. 4. (S.J. O´Brien, ed.); Laboratory of viral carcinogenesis, NIH.

Laemmli, K. (1970):
Cleavage of structural proteins during the assembly of the head of bacteriophage T4.
Nature (London) 227: 680-685

Lang, B.F. (1984):
The mitochondrial genome of the fission yeast Schizosaccaromyces pombe: highly homologous introns are inserted at the same position of the otherwise less conserved cox1 genes in Schizosaccharomyces pombe and Aspergillus nidulans.
EMBO J. 3: 2129-2136

Le Master, D.M. (1986):
Nucleotide sequence and protein overproduction of bacteriophage T4 Thioredoxin.
J. Virol. 59: 759-760

Leong, J.M., S.E. Nunes-Düby, C.F. Lesser, P. Youderian, M.M. Susskind, A. Landy (1985):
Primary structure of the Ø80 and P22 attachment sites in direct interactions with E. coli integration factor.
J. Biol. Chem. 260: 4468-4477

Leong, J.M., S.E. Nunes-Düby, C.F. Lesser, P. Youderian, M.M. Susskind, A. Landy (1986):
Structural and regulatory divergence among site specific recombination genes of lambdoid phages.
J. Mol. Biol. 189: 603-616

Liebig, H.-D. (1987):
Die Konstruktion eines M13-Vektors zur Klonierung und Doppelstrangsequenzierung starker Promotoren: Die Consensussequenz "früher" Promotoren des Bakteriophagen T4.
Dissertation an der Abteilung für Biologie der Ruhr-Universität Bochum

Lilley, D.M. & Kemper, B. (1984):
Cruciform-resolve interactions in supercoiled DNA.
Cell 36: 413-422

Lipman, D.J. & W.R. Pearson (1985):
Rapid and sensitive protein similarity searches.
Science 227: 1435-1441

Luftig, R.B., Wood, W.B., Okinaka, R. (1972):
Bacteriophage T4 head morphogenesis. On the nature of gene 49-defective heads and their role as intermediates.
J. Mol. Biol. 57: 555-573

Macdonald, P.M. & G. Mosig (1984):
Regulation of a new bacteriophage T4 gene, 69, that spans an origin of DNA replication.
EMBO J. 3: 2863-2871

Macdonald, P.M., E. Kutter, G. Mosig (1984):
Regulation of a bacteriophage T4 late gene, soc, which maps in an early region.
Genetics, 106: 17-24

Mattson, T., van Houwe, G., Bolle, A., Selzer, G., Epstein, R. (1977):
Genetic identification of cloned fragments of bacteriophage T4 DNA and complementation by some clones containing early T4 genes.
Mol. Gen. Genet. 156: 319-326

Mc Entee, K., G.M. Weinstock, I.R. Lehmann (1980):
RecA protein-catalyzed strand assimilation: Stimulation by Escherichia coli single stranded DNA binding protein.
Proc. Natl. Acad. Sci. USA 77: 857-861

Mc Donnell, M.W., Simon, M.N., Studier, F.W. (1977):
Analysis of restriction fragments of T7 DNA and determination of molecular-weights by electrophoresis in neutral and alkaline gels.
J. Mol. Biol. 110: 119-146

Mc Pheeters, D.S., A. Christensen, E.T. Young, G. Stormo, L. Gold (1986):
Translational regulation of expression of the bacteriophage T4 lysozyme gene.
Nucl. Acids Res. 14: 5813-5826

Meng, M. & Hogenkamp, H.P.C. (1981):
Purification, characterization, and amino acid sequence of Thioredoxin from Corynebacterium nephridii.
J. Biol. Chem. 256: 9174-9182

Michel, F. & Cummings, D.J. (1985):
Analysis of class I introns in a mitochondrial plasmid associated with senescence of Podospora anserina reveals extraordinary resemblance to the Tetrahymena ribosomal Intron.
Curr. Genet. 10 : 227-239

Michel, F. & B. Dujon (1986):
Genetic Exchanges between Bacteriophage T4 and Filamentous Funghi?
Cell 46: 323

Miller, J., McLachlan, A.D., Klug, A. (1985):
Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes.
EMBO J. 4: 1609-1614

Minagawa, T., A. Murakami, Ryo, Y., H. Yamagishi (1983):
Structural features of very-fast-sedimenting DNA formed by gene 49-defective T4.
Virology 126: 183-193

Mizuuchi, K., Kemper, B., J. Hays, R. Weisberg (1982):
T4 endonuclease VII cleaves Holliday-structures.
Cell 29: 357-365

Morand, P., M. Blanco, R. Devoret (1977):
Characterization of lexB mutations in Escherichia coli K12.
J. Bacteriol. 131: 572-582

Mosig, G. (1983):
Relationship of T4 DNA Replication and Recombination.
In: The Bacteriophage T4, (eds. Mathews, C.K., Kutter, E.M., Mosig, G., Berget, P.B.), pp 120-130, American Society for Microbiology, Washingto, D.C.

Mosig, G. & Macdonald, P.M. (1986):
A new membrane-associated DNA replication protein, the gene 69 product of bacteriophage T4, shares a patch of homology with the Escherichia coli dnaA protein.
J. Mol. Biol. 189: 243-248

Nakabeppu, Y. & M. Sekiguchi (1981):
Physical association of pyrimidine dimer DNA glucosylase and apurinic/apyrimidinic DNA endonuclease essential for repair of ultraviolett-damaged DNA:
Proc. Natl. Acad. Sci. USA 78: 2742-2746

Nash, H.A. (1981):
Integration and excision of bacteriophage λ: The mechanism of conservative site-specific recombination.
Annu. Rev. Genet. 15: 143-167

Nash, H.A. & C.A. Robertson (1981):
Purification and properties of the E. coli protein factor required for λ integrative recombination.
J. Biol. Chem. 256: 9246-9253

Newman, A.K., Rubin, R.A., Kim, s.-H., Modrich, P. (1981):
DNA sequences of structural genes for Eco RI DNA restriction and modification enzymes.
J. Biol. Chem. 256: 2131-2139

Nielsen, H. & J. Engberg (1985):
Sequence comparison of the rDNA introns from six different species of Tetrahymena.
Nucl. Acids Res. 13: 7445-7455

Nishimoto, H., M. Takayama, T. Minagawa (1979):
Purification and some properties of deoxyribonuclease whose synthesis is controlled by gene 49 of bacteriophage T4.
Eur. J. Biochem. 100: 433-440

O´Farrell, P.H., Kutter, E., Nakanishi, M. (1980):
A restriction map of the bacteriophage T4 genome.
Mol. Gen. Genet. 179: 421-435

Pedersen-Lane, J. & M. Belfort (1987):
Variable occurence of the nrdB Intron in the T-even phages suggests Intron mobility.
Science 237: 182-183

Ponticelli, A.S., D.W. Schultz, A.F. Taylor, G.R. Smith (1985):
Chi-dependent DNA strand cleavage by recBC enzyme.
Cell 41: 145-151

Purohit, S. & C.K. Mathews (1984):
Nucleotide sequence reveals overlap between T4 phage genes encoding dihydrofolate reductase and thymidylate synthase.
J. Biol. Chem. 259: 6261-6266

Roberts, J.W. & C.W. Roberts (1981):
Two mutations that alter the regulatory activity of E. coli recA protein.
Nature 290: 422-424

Rosenberg, J.M., J.A. Mc Clarin, C.A. Frederick, B.C. Wang, J. Grable, H.W. Boyer, P. Greene (1987):
Structure and recognition mechanism of Eco RI endonuclease.
TIBS 12: 395-398

Rüger, W. & Kutter, E. (1984):
Bacteriophage T4 genomic map.
In: Genetic maps 1984, O´Brien, S.J. (ed.), Vol. 3, Cold Spring Harbor laboratory press, N.Y. pp27-34

Sancar, A., C. Stachelek, W. Konigsberg, W.D. Rupp (1980):
Sequences of the recA gene and protein.
Proc. Natl. Acad. Sci. USA 77: 2611-2615

Sanger, F., Nicklen, S., Coulson, A.R. (1977):
DNA sequencing with chain-terminating inhibitors.
Proc. Natl. Acad. Sci. USA 5463-5467

Schoemaker, J. (1983):
A T4 function which stabilizes proteins in E. coli.
Trends in Biotech. 1: 99-100

Schumperli, D., McKenney, K., Jobieski, D.A., Rosenberg, M. (1982):
Translational coupling at an intercistronic boundary of the Escherichia coli galactose operon.
Cell 30; 865-871

Schwarz, E., Scherer, G., G. Hobom, Kossel, H. (1978):
Nucleotide sequence of cro, cII, and part of the O gene in phageλ DNA.
Nature 272: 410-414

Shibita, T., C. DasGupta, R.P. Cunningham, C.M. Radding (1980):
Homologous pairing in genetic recombination: Formation of D-loops by combined action of recA protein and a helix destabilizing protein.
Proc. Natl. Acad. Sci. USA 77: 2606-2610

Shine, J. & L. Dalgarno (1974):
The 3´-terminal sequence of Escherichia coli 16s ribosomal RNA: complementary to nonsense triplets and ribosome binding sites.
Proc. Natl. Acad. Sci. USA 71: 1342-1346

Shub, D.A., J.M. Gott, M.-Q. Xu, B.F. Lang, F. Michel, J. Tomaschewski, J. Pedersen-Lane, M. Belfort (1988):
Structural conservation between three homologous Introns of Phage T4 and the group I Introns of eukaryotes.
Proc. Natl. Acad. Sci. USA 85: 1151-1155

Simon, L.D., K. Tomczak, A.C. St. John (1978):
Bacteriophages inhibit degradation of abnormal proteins in E. coli.
Nature (London) 275: 424-428

Simon, L.D. & R.B. Fay (1982):
T4 DNA Fragment as a stabilizer for proteins expressed by cloned DNA.
European Patent Appl. 82106599.2

Sjöberg, B.M. & Holmgren, A. (1972):
Studies on the structure of T4 Thioredoxin: amino acid sequence of the protein and comparison with Thioredoxin of Escherichia coli.
J. Biol. Chem. 247: 8063-8068

Sjöberg, B.M., Hahne, S., Mathews, C.Z., Mathews, C.K., Rand, K.N., Gait, M.J. (1986):
The bacteriophage T4 gene for the small subunit of ribonucleotide reductase contains an Intron.
EMBO J. 5: 2031-2036

Skorupski, K., J. Tomaschewski, W. Rüger, L.D. Simon (1988):
A bacteriophage T4 gene which functions to inhibit Escherichia coli lon protease.
J. Bacteriol., zur Publikation eingereicht

Smith, G.R., S.K. Amundsen, A.M. Chaudhury, K.C. Cheng, A.S. Ponticelli, C.M. Roberts, D.W. Schultz, A.F. Taylor (1984):
Roles of recBC enzyme and chi sites in homologous recombination.
Cold Spring Harbor Symp. Quant. Biol. 49: 485-495

Southern, E.M. (1975):
Detection of specific sequences among DNA fragments separated by gel electrophoresis.
J. Mol. Biol. 98: 503-517

Staden, R. (1980):
A new computer method for the storage and manipulation of DNA gel reading data.
Nucl. Acids Res. 8: 3673-3694

Stahl, F.W. (1979):
Special sites in generalized recombination.
Annu. Rev. Genet. 13: 7-24

Stormo, G.D., Schneider, T.D., Gold, L.M. (1982):
Characterization of translational initiation sites in E. coli.
Nucl. Acids Res. 10: 2971-2996

Studier, F.W. & B.A. Moffatt (1986):
Use of bacteriophage T7 RNA Polymerase to direct selective high level expression of cloned genes.
J. Mol. Biol. 189: 113-130

Symington, L.S., P.T. Morrison, R. Kolodner (1984):
Genetic recombination catalyzed by cell-free extracts of Saccharomyces cerevisiae.
In: Recombination at the DNA-level, (A.J.S. Klar, J.N. Strathern, eds.) Cold Spring Harbor Laboratory N.Y. (USA), pp 803-814

Tabor, S. & C.C. Richardson (1985):
A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes.
Proc. Natl. Acad. Sci. USA 82: 1074-1078

Taylor, A.F. & G.R. Smith (1985):
Substrate specifity of the DNA unwinding activity of the recBC enzyme of Escherichia coli.
J. Mol. Biol. 185: 431-444

Taylor, A.F., D.W. Schultz, A.S. Ponticelli, G.R. Smith (1985):
RecBC enzyme nicking at chi sites during DNA unwinding: location and orientation dependence of the cutting.
Cell, 41: 145-151

Telender-Muskawitz, K.M. & S. Linn (1982):
A unified mechanism for the nuclease and unwinding activities of the recBC enzyme of Escherichia coli.
J. Biol. Chem. 2641-2648

Tessman, I. & D.B. Greenberg (1972):
Ribonucleotide reductase genes of phage T4: map location of the Thioredoxin gene nrdC.
Virology 49: 337-338

Thelander, L. & P. Reichard, P. (1984):
Reduction of Ribonucleotides.
Annu. Rev. Biochem. 48: 133-158

Thomas, B. (1974):
Untersuchungen an der DNA-abhängigen RNA-Polymerase aus Serratia marcescens.
Diplomarbeit an der Abteilung für Biologie der Ruhr-Universität Bochum

Valerie, K., Stevens, J., Lynch, M., Henderson, E.E., de Riel, J.K. (1986):
Nucleotide sequence and analysis of the 58,3 to 65,5 kb early region of bacteriophage T4.
Nucleic Acids Res. 14: 8637-8654

Wahl, G.M., Stern, M., Stark, G.R. (1979):
Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization by using dextran sulfate.
Proc. Natl. Acad. Sci. USA 76: 3683-3687

Wallace, B.J. & S.R. Kushner (1984):
Genetic and physical analysis of the thioredoxin (trxA) gene of Escherichia coli K12.
Gene 32: 399-408

Waring, R.B. & R.W. Davies (1984):
Assessment of a model for Intron RNA secondary structure relevant to RNA self-splicing - a review.
Gene 28: 277-291

Warner, H.R. & D.P. Snustad (1983):
T4 DNA Nucleases.
In: The bacteriophage T4, (Mathews, C.K., Kutter, E.M., Mosig, G., Berget, P.B., eds.) pp 103-109, American Society for Microbiology, Washington, D.C.

Weinberger, C., S.M. Hollenberg, M.G. Rosenfeld, R.M. Evans (1985):
Domain structure of human glucocorticoid receptor and its relationship to the v-erb-A oncogene product.
Nature (London) 318: 670-672

Weisberg, R. & A. Landy (1983):
Site-specific Recombination in Phage Lambda.
In: LambdaII, (eds. Hendrix, R., Roberts, J., Stahl, F., Weisberg, R.), Cold Spring Harbor Laboratory, New York, pp 211-250

Wood, W.B. & H.R. Revel (1976):
The genome of bacteriophage T4.
Bacteriol. Rev. 40: 847-868

Yin, , S., W. Bushman, A. Landy (1985):
Interaction of the site-specific recombination protein xis with attachment site DNA.
Proc. Natl. Acad. Sci. USA 82: 1040-1044

Zuker, M. & Ziegler (1981):
Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information.
Nucl. Acids Res. 9: 133-148

7. ANHANG

Abb. 34: Die Nukleotidsequenz der 9,1 kb-Region.

Zusätzlich sind die potentiellen Promotorstrukturen (____) und die Sekundärstrukturen (---) eingezeichnet.

Abb. 35: Die Nukleotidsequenz des gespleißten sunY-Gens und die hiervon abgeleitete Aminosäuresequenz.

Die Ribosomenbindestelle und das ATG-Initiationskodon sind unterstrichen. Der Pfeil gibt die Position der Spleißstelle an.

Abb. 36: Die P, Q, R und S Sequenzen von nrdB, td und sunY (aus Shub et al., 1988).

Die Homologien zu der jeweiligen Konsensussequenz der Gruppe I und Gruppe IA (s. Burke et al. 1988) sind schattiert dargestellt. Die an der Basenpaarung beteiligten Nukleotide sind durch einen Pfeil gekennzeichnet.

Abb. 37: Sekundärstruktur-Modelle für die Introns von sunY, td und nrdB.

Die Pfeile zeigen die Position der jeweiligen Spleißstelle an. Die Initiations- und Stopkodons der Intron-kodierten Leserahmen sind mit (***) kenntlich gemacht (aus Shub et al., 1988).

Tab. 11: Restriktionskarte für die 9,1 kb-Sequenz.

Insgesamt wurden die Erkennungssequenzen von 95 Restriktionsenzymen berücksichtigt.

Tab. 12: Vergleich des Kodongebrauchs von stark exprimierten *E. coli*-Genen mit der Kodonhäufigkeit von den innerhalb der 9,1 kb-Region angeordneten 19 Genstrukturen.

Tab. 13: Die Kodonhäufigkeit der Gene nrdC, pinA, 49, sunY, 55 und orf 55.12 .

Die unterstrichenen Kodons sind komplementär zu den Antikodons der von T4 kodierten tRNA´s.

Tab. 14: Verzeichnis der in dieser Arbeit konstruierten rekombinanten Vektoren.

```
                                  Y  M  E  T  A  E  K  A  V  R  E  L  A  I  A  Y  Y  M  E  H  G  K  F  P  D  R  Y  S  V  L  K  S  A
                                AATATATGGAGACTGCTGAAAAGGCAGTCCGTGAATTAGCAATAGCTTATTATAATGAACACGGTAAATTTCCTGATAGATACAGCGTGCTTAAATCTGC
        10        20        30        40        50        60        70        80        90       100       110       120

 L  T  R  S  Y  K  N  M  L  S  E  V  S  D  I  I  Y  K  H  K  E  Q  T  G  Q  S  L  D  Y  D  E  T  F  K  Q  V  L  G  I  K
TTTAACTCGTTCATATAAAAATATGCTATCAGAAGTAAGTGATATTATATACAAGCATAAAGAACAAACGGGCCAAAGTCTTGATTACGACGAGACTTTTAAACAAGTACTAGGAATTAA  Pm-1
       130       140       150       160       170       180       190       200       210       220       230       240
         GENE NRDC
 E  *  M  F  K  V  Y  G  Y  D  S  N  I  H  K  C  V  Y  C  D  N  A  K  R  L  L  T  V  K  K  Q  P  F  E  F  I  N  I  M  P
GGAATAATATGTTTAAAGTATATGGTTATGATAGCAACATTCATAAATGTGTGTATTGCGATAATGCAAAACGTCTTTTGACTGTGAAGAAACAGCCGTTTGAATTTATCAACATTATGC
       250       260       270       280       290       300       310       320       330       340       350       360

  E  K  G  V  F  D  D  E  K  I  A  E  L  L  T  K  L  G  R  D  T  Q  I  G  L  T  M  P  Q  V  F  A  P  D  G  S  H  I  G  G
CGGAAAAAGGTGTTTTTGATGATGAGAAAATTGCTGAGCTTCTGACTAAACTAGGTCGTGATACTCAAATCGGCTTAACAATGCCTCAGGTATTTGCTCCTGATGGAAGTCATATTGGTG
       370       380       390       400       410       420       430       440       450       460       470       480
                      ORF 49.3  M  I  E  L  N  E  Q  I  I  F  L  G  D  G  T  E  G  D  L  E  Y  K  L  Y  E  Y  M  I  W  L  A
  F  D  Q  L  R  E  Y  F  K  *
GATTTGACCAATTGCGGGAATACTTTAAATGATAGAATTAAATGAACAAATTATTTTTCTAGGCGATGGAACAGAAGGCGACCTGGAATATAAACTCTATGAATACATGATTTGGTTAGC
       490       500       510       520       530       540       550       560       570       580       590       600

 K  A  E  G  I  D  F  V  V  S  N  P  Y  G  E  N  T  V  V  I  G  G  T  A  Y  E  V  E  W  R  Y  V  G  L  K  S  E  E  Y  D
TAAAGCAGAAGGTATAGACTTTGTTGTGTCTAACCCGTATGGCGAGAACACAGTTGTAATCGGTGGTACCGCTTATGAAGTCGAATGGCGATACGTTGGGCTTAAATCCGAAGAATATGA
       610       620       630       640       650       660       670       680       690       700       710       720
                                                                                ORF 49.2  M  N  I  E  N  K  L  D  V
 V  T  D  E  G  K  W  I  P  I  G  P  W  F  W  E  N  G  E  P  D  F  E  V  S  S  W  W  C  E  K  *
CGTCACCGATGAAGGTAAGTGGATTCCTATCGGCCCCTGGTTCTGGGAACATGGCGAACCCGATTTTGAAGTAAGCAGTTGGTGGTGTGAAAAATGAATATTGAAAATAAATTAGATGTT
       730       740       750       760       770       780       790       800       810       820       830       840

 D  A  V  L  S  E  I  I  E  D  H  D  A  F  S  E  N  Y  D  F  D  F  S  D  Y  L  K  P  I  E  I  E  D  W  V  Q  D  G  K  C
GATGCAGTTCTGAGCGAAATCATTGAAGACCATGATGCATTTTCGGAAAACTATGATTTCGATTTTTCCGACTATCTTAAACCTATCGAAATAGAAGACTGGGTGCAAGATGGTAAATGC
       850       860       870       880       890       900       910       920       930       940       950       960

 Q  Y  R  Q  C  V  Y  F  S  P  K  H  N  V  H  V  A  V  N  E  S  R  S  G  S  Y  H  S  D  W  Y  Y  A  V  P  T  V  E  L  V
CAATATCGTCAGTGTGTTTATTTTAGTCCAAAACATAATGTACATGTGGCTGTAAATGAATCTCGCTCAGGTTCTTACCATTCTGACTGGTATTATGCAGTTCCTACTGTTGAACTGGTT
       970       980       990      1000      1010      1020      1030      1040      1050      1060      1070      1080

Sk-o E  L  R  E  R  V  V  T  Q  T  V  R  E  W  I  T  L  *
GAGCTTCGAGAACGTGTAGTAACTCAGACAGTACGAGAATGGATTACGCTATAAAACCTTGGTGGGCGGCTAGATGGGAAACTGTAGAGCCAGAGCCGGAAGAACCGGTTTACACTGATG
      1090      1100      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
                                                ORF 49.1  M  I  T  V  D  K  W  F  R  I  M  R  A  D  T  G  L  C  N  Y  W  Pe-1
AAGAAACGGTGTATAATGAACCAACGATAAATGACTTAATTGATATGGAGATGGGACATGATTACAGTAGATAAATGGTTTAGAATTAATCGTGCTGATACAGGGCTGTGTAATTACTGG
      1210      1220      1230      1240      1250      1260      1270      1280      1290      1300      1310      1320

 P  E  L  S  A  G  T  V  F  K  V  R  E  L  V  K  E  C  E  D  D  I  E  P  D  T  G  I  I  E  I  E  L  S  D  G  K  I  I  N
CCGGAACTTAGTGCAGGTACTGTCTTTAAAGTTCGTGAACTTGTAAAAGAATGTGAAGATGATATAGAGCCTGACACTGGAATTATTGAAATTGAACTTTCTGATGGAAAGATTATTAAC
      1330      1340      1350      1360      1370      1380      1390      1400      1410      1420      1430      1440

 I  Y  D  K  P  I  T  Y  W  C  L  W  N  T  E  S  V  E  N  G  E  I  E  E  V  V  E  R  T  N  Q  V  V  Q  K  P  K  A  D  F
ATCTACGATAAGCCAATTACGTATTGGTGTTTGTGGAATACTGAATCAGTAGAAAATGGCGAAATTGAAGAAGTTGTAGAGCGAACTAATCAAGTTGTTCAGAAGCCTAAAGCCGATTTT
      1450      1460      1470      1480      1490      1500      1510      1520      1530      1540      1550      1560

 Q  G  E  R  I  S  Y  A  L  A  K  L  A  A  Q  E  N  N  D  G  Y  E  G  N  L  N  Q  A  A  A  E  Y  I  E  W  L  E  T  Q  I
CAAGGTGAACGTATTTCATACGCATTAGCTAAATTAGCTGCACAAGAAAATAACGATGGCTATGAAGGAAATTTAATGCAAGCTGCCGCAGAGTACATTGAATGGCTTGAAACTCAAATT
      1570      1580      1590      1600      1610      1620      1630      1640      1650      1660      1670      1680

Sk-1 S  F  S  D  R  M  I  Q  Q  Y  K  R  L  N  Q  M  F  Y  N  T  *                                    GENE 49  M  L  L  T  G  K
TCTTTTTCTGACCGAATGATTCAGCAATATAAGCGATTGCATCAAATGTTTTACAATACTTGAAAATAATAAATACCCTTATCTATTTAAGGTAAGGGTTTATTATGTTATTGACTGGCA  Pl-1
      1690      1700      1710      1720      1730      1740      1750      1760      1770      1780      1790      1800

  L  Y  K  E  E  K  Q  K  F  Y  D  A  Q  N  G  K  C  L  I  C  Q  R  E  L  N  P  D  V  Q  A  N  H  L  D  H  D  H  E  L  N
AATTATACAAAGAAGAAAAAACAGAAATTTTATGATGCACAAAACGGTAAATGCTTAATTTGCCAACGAGAACTAAATCCTGATGTTCAAGCTAATCACCTCGACCATGACCATGAATTAA
      1810      1820      1830      1840      1850      1860      1870      1880      1890      1900      1910      1920
```

```
 G P K A G K V R G L L C N L C N A A E G Q M K H K F N R S G L K G Q G V D Y L E
ATGGACCAAAAGCAGGAAAGGTGCGTGGATTGCTTTGTAATCTATGCAATGCTGCAGAAGGTCAAATGAAGCATAAATTTAATCGTTCTGGCTTAAAGGGACAAGGTGTTGATTATCTTG
    1930      1940      1950      1960      1970      1980      1990      2000      2010      2020      2030      2040

 W L E N L L T Y L K S D Y T Q N N I H P N F V G D K S K E F S R L G K E E M M A
AATGGTTAGAAAATTTACTTACTTATTTAAAATCCGATTACACCCAAAATAATATTCACCCTAACTTTGTTGGAGATAAATCAAAGGAATTTTCTCGTTTAGGAAAAGAGGAAATGATGG
    2050      2060      2070      2080      2090      2100      2110      2120      2130      2140      2150      2160
                                                                      ORF 55.13  M T I E K E I E G
 E M L Q R G F E Y W E S D T K T Q L I A S F K K Q L R K S L K *
CCGAGATGCTTCAAAGAGGATTTGAATATAATGAATCTGACACCAAAACACAATTAATAGCTTCATTCAAGAAGCAGCTTAGAAAGAGTTTAAAATGACAATTGAAAAAGAAATTGAAGG
    2170      2180      2190      2200      2210      2220      2230      2240      2250      2260      2270      2280  Pm-2

 L I H K T N K D L L N E N A N K D S R V F P T Q R D L M A G I V S K H I A K N M
ATTAATTCATAAAACTAATAAAGACCTTTTAAACGAGAATGCTAATAAAGATTCTCGTGTTTTTCCAACTCAACGGGACCTTATGGCTGGTATTGTGTCTAAACACATTGCCAAAAATAT
    2290      2300      2310      2320      2330      2340      2350      2360      2370      2380      2390      2400

 V P S F I M K A H E S G I I H V H D I D Y S P A L P F T N C C L V D L K G M L E
GGTCCCGTCTTTTATTATGAAAGCGCATGAAAGCGGAATTATTCACGTGCATGATATTGATTATTCCCCTGCTCTTCCATTTACTAATTGCTGTTTAGTAGATTTAAAAGCAATGCTTGA
    2410      2420      2430      2440      2450      2460      2470      2480      2490      2500      2510      2520

 N G F K L G N A Q I E T P K S I G V A T A I M A Q I T A Q V A S H Q Y G G T T F
AAACGGATTTAAGCTTGGTAATGCGCAGATTGAAACTCCTAAATCAATTGGAGTTGCTACTGCAATTATGGCACAAATTACTGCGCAAGTTGCTTCTCACCAATACGGCGGAACGACTTT
    2530      2540      2550      2560      2570      2580      2590      2600      2610      2620      2630      2640

 A N V D K V L S P Y V K R T Y A K H I E D A E K W Q I A D A L N Y A Q S K T E K
TGCGAATGTAGATAAAGTACTTTCTCCTTATGTTAAACGCACCTATGCAAAACATATTGAGGATGCAGAAAAATGGCAAATCGCTGATGCGTTGAATTATGCTCAATCTAAAACAGAAAA
    2650      2660      2670      2680      2690      2700      2710      2720      2730      2740      2750      2760
 D V Y D A F Q A Y E L T *
AGACGTATACGATGCATTCCAAGCTTATGAGTTAACGTAAGTCAAGCTCATGTAAAATCTGCCTAAAACGGGAAACTCTCACTGAGACAATCCGTTGCTAAATCAGCAGTAGCTGTAAAT
    2770      2780      2790      2800      2810      2820      2830      2840      2850      2860      2870      2880

Sk-2 GCCCAACGACTATCCCTGATGAATGTAAGGGAGTAGGGTCAAGCGACCCGAAACGGCAGACAACTCTAAGAGTTGAAGATATAGTCTGAACTGCATGGTGACATGCAGCTGTTTATCCTC
    2890      2900      2910      2920      2930      2940      2950      2960      2970      2980      2990      3000

                      ORF 55.12 M K W K L R K S L K I A N S V A F T Y N V R F P D K S F Y I G
Sk-3 GTATAAATATGAATACGAGGTGAAACGATGAAATGGAAATTAAGAAAAAGCTTAAAAATTGCCAATTCTGTAGCATTTACCTATATGGTAAGATTTCCTGATAAGTCTTTTTATATAGGT
    3010      3020      3030      3040      3050      3060      3070      3080      3090      3100      3110      3120  PI-2

 F K K F K T I Y G K D T N W K E Y N S S S K L V K E K L K D Y K A K W I I L Q V
TTTAAAAAATTCAAAACTATATATGGTAAAGATACAAATTGGAAAGAATACAATTCGTCATCTAAGCTTGTTAAAGAAAAGCTTAAAGATTATAAAGCTAAGTGGATAATTCTTCAAGTT
    3130      3140      3150      3160      3170      3180      3190      3200      3210      3220      3230      3240

 F D S Y E S A L K D E E M L I R K Y F N N E F I L N K S I G G Y K F N K Y P D S
TTTGATTCTTATGAATCGGCCCTTAAAGATGAAGAAATGCTTATTAGGAAATATTTTAATAACGAATTTATTCTTAATAAATCTATAGGTGGATCTAAATTTAACAAATATCCGGATTCA
    3250      3260      3270      3280      3290      3300      3310      3320      3330      3340      3350      3360

 E E M K Q K L S N A H K G K I L S L K H K D K I R E K L I E H Y K N N S R S E A
GAAGAACATAAGCAAAAACTTAGTAATGCCCATAAAGGTAAAATCTTATCTTTAAAACATAAAGATAAGATACGAGAGAAATTGATTGAGCATTATAAAAATAATAGTAGAAGTGAAGCT
    3370      3380      3390      3400      3410      3420      3430      3440      3450      3460      3470      3480

 H V K N N I G S R T A K K T V S I A L K S G N K F R S F K S A A K F L K C S E E
CATGTTAAAAATAATATTGGTAGTAGAACGGCTAAAAAGACTGTTTCTATAGCTTTAAAATCCGGAAATAAATTTAGAAGTTTTAAATCAGCTGCAAAATTTCTTAAATGCTCTGAAGAA
    3490      3500      3510      3520      3530      3540      3550      3560      3570      3580      3590      3600

 Q V S N H P N V I D I K I T I H P V P E Y V K I N D N I Y K S F V D A A K D L K
CAGGTTAGTAATCATCCAAATGTTATAGATATAAAAATAACAATTCATCCCGTCCCAGAATATGTTAAAATAAATGACAATATCTATAAATCATTTGTGGATGCTGCTAAAGATTTAAAA
    3610      3620      3630      3640      3650      3660      3670      3680      3690      3700      3710      3720

 L H P S R I K D L C L D D N Y P N Y I V S Y K R V E K *
CTCCACCCAAGTCGTATTAAAGATTTGTGTTTAGATGACAATTATCCAAATTACATTGTTTCATATAAACGGGTAGAGAAGTAGCGAACTCTACTGAACACATTGATGAAGTGAACACGT
    3730      3740      3750      3760      3770      3780      3790      3800      3810      3820      3830      3840
```

```
                                                                                ORF 55.11 M  I  Q  K  A  I  L  K  N  R  I  K  G  L  G  R  D
TATTCAGTTCAAACGGACAGACTCCTTTTGTAACAATTACATTTGGTACGGGAACTGACTGGACTGAACGAATGATTCAGAAAGCAATTCTGAAAAATCGTATTAAAGGTCTTGGTCGTG
      3850      3860      3870      3880      3890      3900      3910      3920      3930      3940      3950      3960

  G  I  T  P  I  F  P  K  L  V  M  F  V  E  E  G  V  N  L  Y  K  D  D  P  W  Y  D  I  K  Q  L  A  L  E  C  A  S  K  R  M
ATGGGATAACTCCTATTTTCCCTAAGCTTGTTATGTTCGTTGAAGAAGGTGTTAATCTTTATAAAGACGATCCGAACTATGATATTAAACAGCTTGCTCTAGAGTGCGCAAGCAAAAGGA
      3970      3980      3990      4000      4010      4020      4030      4040      4050      4060      4070      4080

  Y  P  D  I  I  S  A  K  N  N  K  A  I  T  G  S  S  V  P  V  S  P  W  G  C  R  S  F  L  S  V  W  K  D  S  T  G  N  E  I
TGTATCCTGATATTATTTCAGCTAAGAACAATAAAGCTATCACTGGTTCATCTGTTCCTGTTTCTCCGATGGGTTGCCGTAGTTTCTTGAGCGTATGGAAAGATTCGACTGGCAATGAAA
      4090      4100      4110      4120      4130      4140      4150      4160      4170      4180      4190      4200

  L  D  G  R  W  N  L  G  V  V  T  L  N  L  P  R  I  A  L  D  S  Y  I  G  T  Q  F  N  E  Q  K  F  V  E  L  F  W  E  R  M
TTCTTGATGGACGCAATAATCTTGGTGTTGTAACACTGAATCTTCCTCGCATCGCGTTAGATTCTTATATTGGAACACAGTTCAATGAACAGAAATTTGTTGAGCTATTTAATGAACGAA
      4210      4220      4230      4240      4250      4260      4270      4280      4290      4300      4310      4320

  D  L  C  F  E  A  L  M  C  R  I  S  S  L  K  G  V  K  A  T  V  A  P  I  L  Y  Q  E  G  A  F  G  V  R  L  K  P  D  D  D
TGGATTTATGTTTTGAAGCTTTGATGTGTAGAATTAGTTCCTTAAAAGGAGTTAAAGCTACTGTTGCTCCTATTCTTTACCAAGAAGGTGCATTCGGGGTTCGTCTTAAACCTGATGACG
      4330      4340      4350      4360      4370      4380      4390      4400      4410      4420      4430      4440

  I  I  E  L  F  K  N  G  R  S  S  V  S  L  G  Y  I  G  I  H  E  L  N  I  L  V  G  R  D  I  G  R  E  I  L  T  K  M  N  A
ACATAATTGAGTTATTTAAAAACGGTAGAAGTTCAGTGTCTTTAGGATACATTGGTATTCACGAATTGAATATTCTTGTCGGTCGTGATATTGGACGAGAAATTTTAACTAAAATGAATG
      4450      4460      4470      4480      4490      4500      4510      4520      4530      4540      4550      4560

  H  L  K  Q  W  T  E  R  T  G  F  A  F  S  L  Y  S  T  P  A  E  N  L  C  Y  R  F  C  K  L  D  T  E  K  Y  G  S  V  K  D
CTCATCTTAAACAGTGGACTGAAAGAACCGGATTTGCTTTTAGTTTATATTCGACTCCTGCTGAAAACCTGTGTTATCGCTTCTGTAAACTCGATACAGAAAAATATGGAAGCGTAAAAG
      4570      4580      4590      4600      4610      4620      4630      4640      4650      4660      4670      4680

  V  T  D  K  G  W  Y  T  N  S  F  M  V  S  V  E  E  N  I  T  P  F  E  K  I  S  R  E  A  P  Y  H  F  I  A  T  G  G  H  I
ATGTTACCGATAAAGGCTGGTACACTAACAGTTTCCATGTTTCAGTAGAAGAAAATATTACTCCGTTTGAAAAGATTTCTCGTGAAGCGCCATATCATTTCATTGCGACAGGTGGTCACA
      4690      4700      4710      4720      4730      4740      4750      4760      4770      4780      4790      4800

  S  Y  V  E  L  P  D  W  K  N  N  L  K  G  L  E  A  V  W  D  Y  A  A  Q  H  L  D  Y  F  G  V  N  M  P  V  D  K  C  F  T
TTTCTTATGTTGAACTTCCTGATATGAAAAATAACTTAAAAGGTCTTGAGGCCGTGTGGGATTATGCTGCACAGCATTTAGATTATTTTGGTGTTAATATGCCAGTAGATAAATGTTTTA
      4800      4810      4820      4830      4840      4850      4860      4870      4880      4890      4900      4910

  C  G  S  T  H  E  M  T  P  T  E  N  G  F  V  C  S  I  C  G  E  T  D  P  K  K  M  N  T  I  R  R  T  C  G  Y  L  G  N  P
CATGTGGAAGTACCCATGAAATGACTCCTACTGAAAACGGATTTGTTTGTTCTATTTGTGGAGAAACTGATCCTAAAAAGATGAATACCATAAGAAGAACATGTGGTTATTTGGGAAATC
      4930      4940      4950      4960      4970      4980      4990      5000      5010      5020      5030      5040
                                                                ORF 55.10 M  N  Y  D  R  I  Y  S  N  L  V  N  S  A  E  H  P  S  K  P
  W  E  R  G  F  N  L  G  K  N  K  E  I  M  H  R  V  K  H  Q  *
CCAACGAACGCGGATTTAATCTCGGTAAAAATAAAGAAATCATGCATAGGGTTAAGCACCAATGAATTATGACAGAATCTATTCAAATCTAGTTAATTCTGCTGAACACCCATCTAAGCC
      5050      5060      5070      5080      5090      5100      5110      5120      5130      5140      5150      5160

  R  Q  T  K  A  G  Y  E  L  H  H  I  I  P  R  S  M  G  G  S  D  D  L  D  N  L  V  F  L  T  Y  K  A  H  Y  T  A  H  H  L
CCGTCAAACTAAAGCCGGGTATGAACTTCATCATATAATTCCAAGAAGTATGGGTGGTTCAGATGATTTAGACAACTTAGTATTTTTAACTTATAAGGCCCACTATACTGCTCACCATTT
      5170      5180      5190      5200      5210      5220      5230      5240      5250      5260      5270      5280

  L  A  K  I  Y  G  G  K  M  Y  D  A  Y  W  I  M  S  N  S  K  H  R  K  V  T  S  S  Q  Y  A  Y  A  K  A  K  K  F  E  N  Q
ATTGGCTAAAATATATGGTGGAAAGATGTATGATGCATATTGGATAATGAGTAATAGTAAGCATAGAAAAGTTACTTCTTCACAATACGCGTATGCAAAAGCTAAGAAATTTGAAAATCA
      5290      5300      5310      5320      5330      5340      5350      5360      5370      5380      5390      5400

  R  G  A  K  R  S  D  K  T  K  K  L  M  S  I  A  R  S  K  V  T  L  P  D  D  F  G  K  R  I  S  D  G  L  T  G  R  T  L  S
GCGTGGCGCAAAAAGAAGCGATAAAACTAAGAAGCTAATGAGCATAGCTAGGTCTAAAGTAACTCTTCCTGATGATTTTGGAAAAAGGATATCTGATGGGCTAACAGGAAGAACTTTGTC
      5410      5420      5430      5440      5450      5460      5470      5480      5490      5500      5510      5520

  T  N  H  K  E  H  I  K  N  S  S  R  N  R  F  N  T  K  T  V  I  G  Y  S  P  D  G  N  T  I  E  L  T  G  A  T  E  I  R  N
TACTAATCATAAAGAGCATATCAAAAATAGTTCTAGAAATCGCTTTAACACCAAAACGGTTATTGGCTATTCACCAGACGGAAATACAATAGAACTAACTGGAGCAACTGAAATAAGAAA
      5530      5540      5550      5560      5570      5580      5590      5600      5610      5620      5630      5640
                                                                                ORF 55.9  M  N  Y  D  R  I  Y  P  C  D  F  V  N
  A  G  F  F  H  S  A  V  Y  K  C  C  K  G  I  Q  S  I  H  K  G  Y  T  W  K  Y  K  D  E  L  *
TGCAGGATTTTTTCATTCTGCTGTTTATAAATGTTGCAAAGGAATTCAATCTATTCATAAGGGATATACTTGGAAATATAAAGATGAATTATGATAGAATTTATCCATGCGATTTCGTTA
      5650      5660      5670      5680      5690      5700      5710      5720      5730      5740      5750      5760
```

```
Sk-4  G  P  G  C  R  V  V  L  F  V  T  G  C  L  H  K  C  E  G  C  Y  N  R  S  T  W  N  A  R  N  G  Q  L  F  T  M  N  T  V  K
      ACGGCCCTGGATGTAGGGTCGTTCTCTTCGTCACTGGATGTTTGCATAAATGCGAAGGTTGTTATAATCGTAGTACTTGGAATGCTCGTAACGGACAGTTATTCACTATGAATACCGTTA
            5770      5780      5790      5800      5810      5820      5830      5840      5850      5860      5870      5880

        E  L  A  S  H  L  S  K  S  Y  I  Q  G  L  T  L  T  G  G  D  P  L  Y  P  Q  N  R  E  E  I  S  N  L  V  S  W  V  K  A  R
      AAGAACTTGCGTCCCATTTAAGCAAATCGTATATCCAAGGCCTTACCTTAACCGGCGGTGACCCACTTTATCCACAGAACCGAGAAGAGATTTCAAATTTAGTTTCTTGGGTTAAAGCAA
            5890      5900      5910      5920      5930      5940      5950      5960      5970      5980      5990      6000

        F  P  E  K  D  I  W  L  W  T  G  Y  K  F  E  D  I  K  Q  L  E  M  L  K  Y  V  D  V  I  I  D  G  K  Y  E  K  N  L  P  T
      GATTTCCGGAGAAAGATATCTGGTTGTGGACAGGATATAAGTTTGAAGATATTAAACAACTAGAAATGCTTAAATATGTTGATGTTATTATTGATGGGAAGTATGAGAAAAATCTTCCGA
            6010      6020      6030      6040      6050      6060      6070      6080      6090      6100      6110      6120  Pm-3

        K  K  L  W  R  G  S  D  N  Q  R  L  W  S  N  T  D  G  V  W  K  H  D  *
      CCAAAAAGCTGTGGCGAGGATCAGATAATCAGCGACTTTGGTCAAATACCGATGGGGTGTGGAAACATGATTAACTTGAATTACATTATGGATACTATAAATGATATGATTTTCATTTTG
            6130      6140      6150      6160      6170      6180      6190      6200      6210      6220      6230      6240

                                                                ORF 55.8  M  Y  K  F  R  K  G  L  A  D  F  L  T  T  V  T  F
      GTCCAGAATTTTATTCGCAGTATAGTTTAGTGCTTATCAATGCTTGGTTAATCAATTAAAGGTAAATGTAAATGTATAAATTTCGTAAAGGTTTAGCTGATTTTCTTACAACTGTAACAT
            6250      6260      6270      6280      6290      6300      6310      6320      6330      6340      6350      6360

        F  L  F  M  A  V  G  A  I  F  L  I  P  F  I  A  I  F  F  V  I  S  L  I  S  P  E  K  G  L  S  S  S  E  F  N  E  R  L  D
      TCTTTCTGTTTATGGCAGTTGGAGCTATTTTCCTTATTCCTTTTATTGCTATATTTTTCGTGATTAGTTTGATTTCTCCAGAAAAGGGCTTATCTTCTAGTGAATTCAATGAGCGTCTGG
            6370      6380      6390      6400      6410      6420      6430      6440      6450      6460      6470      6480
                                                      ORF 55.7
        K  I  T  N  K  L  N  A  A  L  S  K  E  *   M  K  Q  N  K  I  E  V  Y  G  I  P  D  E  V  G  R  C  P  G  C  Q  S  V  T
      ATAAAATTACTAACAAGCTGAATGCTGCTCTTAGTAAGGAATAGTTGTGAAACAAAATAAGATTGAAGTCTATGGAATTCCAGATGAAGTAGGTCGTTGTCCTGGATGTCAATCAGTTAC
            6490      6500      6510      6520      6530      6540      6550      6560      6570      6580      6590      6600

        K  L  L  K  E  L  N  A  P  F  T  F  Y  K  V  L  T  N  N  G  K  I  E  Y  D  R  P  L  I  V  S  L  A  K  R  A  G  F  T  S
      AAAACTTCTAAAGGAGCTCAATGCTCCTTTTACTTTCTATAAAGTTCTTACAAATAATGGTAAGATTGAGTATGATCGTCCGCTGATTGTATCTCTTGCTAAACGCGCTGGATTCACATC
            6610      6620      6630      6640      6650      6660      6670      6680      6690      6700      6710      6720

        L  N  I  R  Y  P  V  I  F  I  N  D  S  R  Q  K  N  I  K  H  F  K  E  T  L  I  S  L  G  Y  D  R  D  I  I  E  D  *
      TCTTAACATTCGTTATCCAGTCATTTTCATTAATGATTCTAGACAAAAGAACATTAAACACTTCAAAGAAACTCTCATTTCACTTGGATATGATAGAGATATCATAGAAGATTAAGACGG
            6730      6740      6750      6760      6770      6780      6790      6800      6810      6820      6830      6840

Sk-5  GCCCTCTGGGCCTTTCTTTCTCACATTCTGTATATTACCATTCTAAGCTATCGTTCCTTTCTTATCATTCCCTAAAATATTTTTCACAAAGTTGTTTACAACAAGTTCAAACCGTGGTAT
            6850      6860      6870      6880      6890      6900      6910      6920      6930      6940      6950      6960
                                                                                                                              Pe-2
                                ORF 55.6  M  V  V  V  D  K  E  I  K  K  G  Q  Y  Y  L  V  N  G  N  V  V  R  V  T  Y  V  N  G  F
      TATTAACATATGAATTACCTTTGAGGAATTGATATGGTTGTGGTTGATAAAGAGATTAAAAAGGGACAGTATTATCTCGTGAATGGTAATGTTGTTCGTGTTACTTATGTAAATGGTTTT
            6970      6980      6990      7000      7010      7020      7030      7040      7050      7060      7070      7080
                                                                                ORF 55.5  M  G  K  T  Y  R  R  K  D  L  K
      D  V  Y  Y  L  I  L  K  L  H  K  R  M  I  C  D  R  A  V  F  S  S  V  A  K  E  I  K  L  H  G  *
      GATGTTTACTATCTTATACTCAAGTTACATAAACGTATGATTTGTGATCGTGCTGTATTTAGTTCAGTTGCTAAGGAAATTAAACTCCATGGGTAAAACGTATCGTCGTAAAGACTTAAA
            7090      7100      7110      7120      7130      7140      7150      7160      7170      7180      7190      7200

        V  R  D  Y  D  Y  F  G  K  R  K  A  P  D  G  V  S  H  K  D  M  V  E  N  I  F  R  S  D  K  W  R  R  M  K  G  I  D  S  E
      AGTGCGTGATTATGACTATTTCGGAAAGCGTAAAGCTCCTGATGGGGTAAGTCATAAAGCTATGGTTGAAAACATTTTTCGCTCAGATAAATGGCGTAGAATGAAAGGTATTGATTCAGA
            7210      7220      7230      7240      7250      7260      7270      7280      7290      7300      7310      7320

        V  K  D  E  L  N  R  Q  L  R  G  E  V  R  K  L  K  K  S  V  Y  I  D  D  D  F  D  Y  N  T  S  Q  R  V  A  K  R  K  S  N
      AGTTAAAGATGAGTTAAATCGTCAATTACGCGGTGAAGTAAGAAAGTTGAAAAAATCAGTTTACATTGATGATGATTTTGATTATAATACTTCTCAACGAGTTGCTAAACGTAAATCAAA
            7330      7340      7350      7360      7370      7380      7390      7400      7410      7420      7430      7440  Pe-3
                                ORF 55.4
        E  C  Y  R  Y  S  *     M  N  I  K  R  M  L  F  K  Q  G  L  Y  T  L  N  V  T  P  K  G  D  T  T  K  W  S  V  N  D  W
      CGAGTGTTATCGTTACAGCTGAGGAAAATATGAATATCAAACGCATGCTTTTTAAGCAAGGGCTATACACTTTAAATGTTACTCCAAAAGGCGATACGACCAAGTGGTCAGTAAATGACT
            7450      7460      7470      7480      7490      7500      7510      7520      7530      7540      7550      7560
                                                      ORF 55.3
        I  K  F  I  D  E  N  G  N  W  E  I  *  M  N  P  E  S  K  L  S  Q  R  I  A  E  E  R  A  K  F  F  Q  N  M  K  H  N  G  I
      GGATTAAATTTATTGATGAAAATGGCAATTGGGAAATTTAAATGAATCCTGAATCTAAATTATCGCAGGCAATTGCTGAAGAACGCGCCAAATTTTTCCAGAACATGAAACACAACGGTA
            7570      7580      7590      7600      7610      7620      7630      7640      7650      7660      7670      7680
```

```
 E  D  E  V  F  L  N  W  F  W  N  N  K  Y  A  A  C  E  G  A  L  S  L  S  V  A  N  M  Y  E  G  W  K  G  A  K  K  F  S  L
TTGAGGATGAAGTTTTTCTAAATTGGTTCTGGAATAATAAGTATGCAGCATGTGAAGGAGCTTTGTCATTGTCAGTCGCAATGATGTACGAAGGCTGGAAGGGTGCCAAAAAGTTTAGCC
    7690      7700      7710      7720      7730      7740      7750      7760      7770      7780      7790      7800

 R  A  S  A  F  L  D  N  K  I  L  T  *     ORF 55.2  M  T  I  Q  I  K  N  A  I  N  S  Y  A  Y  D  K  V  V  S  L  L  E
TAAGGGCTTCGGGCCTTTTTGGATAATAAAATTTTAACGTAATTGAGGATAATGTATGACTATTCAAATTAAAAACGCCATCAATTCTTACGCATATGATAAAGTAGTTTCTTTGCTAGAA
    7810      7820      7830      7840      7850      7860      7870      7880      7890      7900      7910      7920

 K  G  D  I  V  T  P  Q  I  L  D  K  W  E  K  E  L  H  Q  T  M  K  Q  N  D  Q  K  I  G  R  N  T  V  R  E  L  L  V  Q  Y
AAAGGCGATATTGTAACTCCTCAAATTTTGGATAAATGGGAAAAAGAGCTTCATCAAACGATGAAACAGAATGATCAGAAGATTGGACGCAATACTGTCCGTGAATTGTTGGTTCAATAT
    7930      7940      7950      7960      7970      7980      7990      8000      8010      8020      8030      8040

 I  L  S  E  F  D  V  K  A  F  G  V  E  S  K  A  Y  Q  K  H  E  I  S  D  K  T  I  R  R  M  K  N  Q  R  K  K  K  F  A  D
ATCTTGTCAGAATTTGATGTTAAAGCTTTTGGTGTAGAATCTAAAGCTTATCAAAAGCATGAAATTTCCGATAAAACTATTCGTCGCATGAAAAATCAACGCAAGAAAAAATTTGCAGAC
    8050      8060      8070      8080      8090      8100      8110      8120      8130      8140      8150      8160
                        ORF 55.1
 L  K  I  T  K  V  *  M  N  E  A  L  I  N  D  L  R  L  A  G  Y  E  V  N  T  N  G  I  G  L  T  Q  I  E  G  N  G  F  I  L
CTGAAAATTACTAAGGTATAATTATGAACGAAGCTCTTATTAACGATTTACGTCTCGCTGGGTATGAAGTAAATACAAATGGCATTGGTTTAACTCAAATTGAAGGAAATGGATTCATCC
    8170      8180      8190      8200      8210      8220      8230      8240      8250      8260      8270      8280

 E  Y  E  F  S  Q  W  W  L  Y  A  N  Y  G  E  L  I  E  Y  V  D  Q  F  D  S  L  D  A  A  L  G  A  A  N  L  M  K  C  L  K
TTGAGTATGAATTTAGCCAATGGTGGTTATACGCTAATTACGGCGAATTAATTGAATATGTTGACCAATTTGATTCACTAGATGCAGCTCTTGGAGCGGCTAATCTGATGAAGTGCTTGA
    8290      8300      8310      8320      8330      8340      8350      8360      8370      8380      8390      8400

 L  I  H  P  L  C  F  K  I  G  L  Q  I  L  *                          GENE 55  M  S  E  T  K  P  K  Y  N  Y  V  N  N
AATTGATTCATCCTCTATGCTTTAAGATAGGTCTTCAAATATTATGATATAATAGATCTATGAATTGAGCTAAGAGGTGAAAATGTCAGAAACTAAGCCTAAATATAATTACGTAAACAA
    8410      8420      8430      8440      8450      8460      8470      8480      8490      8500      8510      8520

 K  E  L  L  Q  A  I  I  D  W  K  T  E  L  A  M  N  K  D  P  N  K  V  V  R  Q  N  D  T  I  G  L  A  I  M  L  I  A  E  G
TAAAGAGCTTTTACAAGCTATTATTGATTGGAAAACAGAATTAGCAAATAATAAAGACCCAAATAAAGTAGTTCGTCAGAATGATACTATCGGATTAGCCATTATGCTTATTGCAGAAGG
    8530      8540      8550      8560      8570      8580      8590      8600      8610      8620      8630      8640

 L  S  K  R  F  N  F  S  G  Y  T  Q  S  W  K  Q  E  M  I  A  D  G  I  E  A  S  I  K  G  L  H  N  F  D  E  T  K  Y  K  N
CTTATCTAAACGTTTCAACTTTTCAGGATACACCCAGTCTTGGAAACAAGAAATGATTGCCGATGGTATAGAAGCTTCTATTAAGGGGCTTCACAATTTTGATGAAACGAAATATAAAAA
    8650      8660      8670      8680      8690      8700      8710      8720      8730      8740      8750      8760

 P  H  A  Y  I  T  Q  A  C  F  N  A  F  V  Q  R  I  K  K  E  R  K  E  V  A  K  K  Y  S  Y  F  V  H  N  V  Y  D  S  R  D
CCCACATGCGTATATAACTCAAGCTTGTTTTAATGCATTCGTCCAACGTATTAAAAAAGAACGTAAGGAAGTTGCAAAGAAATATAGTTACTTCGTTCACAATGTCTATGACAGCCGTGA
    8770      8780      8790      8800      8810      8820      8830      8840      8850      8860      8870      8880

 D  D  M  V  A  L  V  D  E  T  F  I  Q  D  I  Y  D  K  M  T  N  Y  E  E  S  T  Y  R  T  P  G  A  E  K  K  S  V  V  D  D
CGACGATATGGTTGCGTTAGTAGATGAAACTTTTATTCAAGACATCTATGATAAAATGACGCATTACGAAGAATCAACCTATAGAACACCGGGGGCTGAAAAGAAAAGTGTTGTAGATGA
    8890      8900      8910      8920      8930      8940      8950      8960      8970      8980      8990      9000

 S  S  S  L  D  F  L  Y  E  A  N  D  *
TTCTTCTAGTTTGGATTTTTTATATGAGGCTAACGATTAACCTCTCCGGATTCTTGGAAGAAATACCTGAAGTTGAAGCTATTCCCTATTTACTTAAAATGTATCTCAGGGAAGTTTTAG
    9010      9020      9030      9040      9050      9060      9070      9080      9090      9100      9110      9120
```

Abbildung 34.

```
                                MetThrIleGluLysGluIleGluGlyLeuIleHis
GAAGCAGCTTAGAAAGAGTTTAAAATGACAATTGAAAAAGAAATTGAAGGATTAATTCAT
         10        20        30        40        50        60

LysThrAsnLysAspLeuLeuAsnGIuAsnAlaAsnLysAspSerArgValPheProThr
AAAACTAATAAAGACCTTTTAAACGAGAATGCTAATAAAGATTCTCGTGTTTTTCCAACT
         70        80        90       100       110       120

GlnArgAspLeuMetAlaGlyIleValSerLysHisIleAlaLysAsnMetValProSer
CAACGGGACCTTATGGCTGGTATTGTGTCTAAACACATTGCCAAAAATATGGTCCCGTCT
        130       140       150       160       170       180

PheIleMetLysAlaHisGluSerGlyIleIleHisValHisAspIleAspTyrSerPro
TTTATTATGAAAGCGCATGAAAGCGGAATTATTCACGTGCATGATATTGATTATTCCCCT
        190       200       210       220       230       240

AlaLeuProPheThrAsnCysCysLeuValAspLeuLysGlyMetLeuGluAsnGlyPhe
GCTCTTCCATTTACTAATTGCTGTTTAGTAGATTTAAAAGGAATGCTTGAAAACGGATTT
        250       260       270       280       290       300

LysLeuGlyAsnAlaGlnIleGluThrProLysSerIleGlyValAlaThrAlaIleMet
AAGCTTGGTAATGCGCAGATTGAAACTCCTAAATCAATTGGAGTTGCTACTGCAATTATG
        310       320       330       340       350       360

AlaGlnIleThrAlaGlnValAlaSerHisGlnTyrGlyGlyThrThrPheAlaAsnVal
GCACAAATTACTGCGCAAGTTGCTTCTCACCAATACGGCGGAACGACTTTTGCGAATGTA
        370       380       390       400       410       420

AspLysValLeuSerProTyrValLysArgThrTyrAlaLysHisIleGluAspAlaGlu
GATAAAGTACTTTCTCCTTATGTTAAACGCACCTATGCAAAACATATTGAGGATGCAGAA
        430       440       450       460       470       480

LysTrpGlnIleAlaAspAlaLeuAsnTyrAlaGlnSerLysThrGluLysAspValTyr
AAATGGCAAATCGCTGATGCGTTGAATTATGCTCAATCTAAAACAGAAAAAGACGTATAC
        490       500       510       520       530       540

                          ↓
AspAlaPheGlnAlaTyrGluTyrGluValAsnThrLeuPheSerSerAsnGlyGlnThr
GATGCATTCCAAGCTTATGAGTATGAAGTGAACACGTTATTCAGTTCAAACGGACAGACT
        550       560       570       580       590       600

ProPheValThrIleThrPheGlyThrGlyThrAspTrpThrGluArgMetIleGlnLys
CCTTTTGTAACAATTACATTTGGTACGGGAACTGACTGGACTGAACGAATGATTCAGAAA
        610       620       630       640       650       660

AlaIleLeuLysAsnArgIleLysGlyLeuGlyArgAspGlyIleThrProIlePhePro
GCAATTCTGAAAAATCGTATTAAAGGTCTTGGTCGTGATGGGATAACTCCTATTTTCCCT
        670       680       690       700       710       720

LysLeuValMetPheValGluGluGlyValAsnLeuTyrLysAspAspProAsnTyrAsp
AAGCTTGTTATGTTCGTTGAAGAAGGTGTTAATCTTTATAAAGACGATCCGAACTATGAT
        730       740       750       760       770       780

IleLysGlnLeuAlaLeuGluCysAlaSerLysArgMetTyrProAspIleIleSerAla
ATTAAACAGCTTGCTCTAGAGTGCGCAAGCAAAAGGATGTATCCTGATATTATTTCAGCT
        790       800       810       820       830       840

LysAsnAsnLysAlaIleThrGlySerSerValProValSerProMetGlyCysArgSer
AAGAACAATAAAGCTATCACTGGTTCATCTGTTCCTGTTTCTCCGATGGGTTGCCGTAGT
        850       860       870       880       890       900

PheLeuSerValTrpLysAspSerThrGlyAsnGluIleLeuAspGlyArgAsnAsnLeu
TTCTTGAGCGTATGGAAAGATTCGACTGGCAATGAAATTCTTGATGGACGCAATAATCTT
        910       920       930       940       950       960
```

```
GlyValValThrLeuAsnLeuProArgIleAlaLeuAspSerTyrIleGlyThrGlnPhe
GGTGTTGTAACACTGAATCTTCCTCGCATCGCGTTAGATTCTTATATTGGAACACAGTTC
        970       980       990      1000      1010      1020

AsnGluGlnLysPheValGluLeuPheAsnGluArgMetAspLeuCysPheGluAlaLeu
AATGAACAGAAATTTGTTGAGCTATTTAATGAACGAATGGATTTATGTTTTGAAGCTTTG
       1030      1040      1050      1060      1070      1080

MetCysArgIleSerSerLeuLysGlyValLysAlaThrValAlaProIleLeuTyrGln
ATGTGTAGAATTAGTTCCTTAAAAGGAGTTAAAGCTACTGTTGCTCCTATTCTTTACCAA
       1090      1100      1110      1120      1130      1140

GluGlyAlaPheGlyValArgLeuLysProAspAspAspIleIleGluLeuPheLysAsn
GAAGGTGCATTCGGGGTTCGTCTTAAACCTGATGACGACATAATTGAGTTATTTAAAAAC
       1150      1160      1170      1180      1190      1200

GlyArgSerSerValSerLeuGlyTyrIleGlyIleHisGluLeuAsnIleLeuValGly
GGTAGAAGTTCAGTGTCTTTAGGATACATTGGTATTCACGAATTGAATATTCTTGTCGGT
       1210      1220      1230      1240      1250      1260

ArgAspIleGlyArgGluIleLeuThrLysMetAsnAlaHisLeuLysGlnTrpThrGlu
CGTGATATTGGACGAGAAATTTTAACTAAAATGAATGCTCATCTTAAACAGTGGACTGAA
       1270      1280      1290      1300      1310      1320

ArgThrGlyPheAlaPheSerLeuTyrSerThrProAlaGluAsnLeuCysTyrArgPhe
AGAACCGGATTTGCTTTTAGTTTATATTCGACTCCTGCTGAAAACCTGTGTTATCGCTTC
       1330      1340      1350      1360      1370      1380

CysLysLeuAspThrGluLysTyrGlySerValLysAspValThrAspLysGlyTrpTyr
TGTAAACTCGATACAGAAAAATATGGAAGCGTAAAAGATGTTACCGATAAAGGCTGGTAC
       1390      1400      1410      1420      1430      1440

ThrAsnSerPheHisValSerValGluGluAsnIleThrProPheGluLysIleSerArg
ACTAACAGTTTCCATGTTTCAGTAGAAGAAAATATTACTCCGTTTGAAAAGATTTCTCGT
       1450      1460      1470      1480      1490      1500

GluAlaProTyrHisPheIleAlaThrGlyGlyHisIleSerTyrValGluLeuProAsp
GAAGCGCCATATCATTTCATTGCGACAGGTGGTCACATTTCTTATGTTGAACTTCCTGAT
       1510      1520      1530      1540      1550      1560

MetLysAsnAsnLeuLysGlyLeuGluAlaValTrpAspTyrAlaAlaGlnHisLeuAsp
ATGAAAAATAACTTAAAAGGTCTTGAGGCCGTGTGGGATTATGCTGCACAGCATTTAGAT
       1570      1580      1590      1600      1610      1620

TyrPheGlyValAsnMetProValAspLysCysPheThrCysGlySerThrHisGluMet
TATTTTGGTGTTAATATGCCAGTAGATAAATGTTTTACATGTGGAAGTACCCATGAAATG
       1630      1640      1650      1660      1670      1680

ThrProThrGluAsnGlyPheValCysSerIleCysGlyGluThrAspProLysLysMet
ACTCCTACTGAAAACGGATTTGTTTGTTCTATTTGTGGAGAAACTGATCCTAAAAAGATG
       1690      1700      1710      1720      1730      1740

AsnThrIleArgArgThrCysGlyTyrLeuGlyAsnProAsnGluArgGlyPheAsnLeu
AATACCATAAGAAGAACATGTGGTTATTTGGGAAATCCGAACGAACGCGGATTTAATCTC
       1750      1760      1770      1780      1790      1800

GlyLysAsnLysGluIleMetHisArgValLysHisGln***
GGTAAAAATAAAGAAATCATGCATAGGGTTAAGCACCAATGA
       1810      1820      1830      1840
```

Abbildung 35.

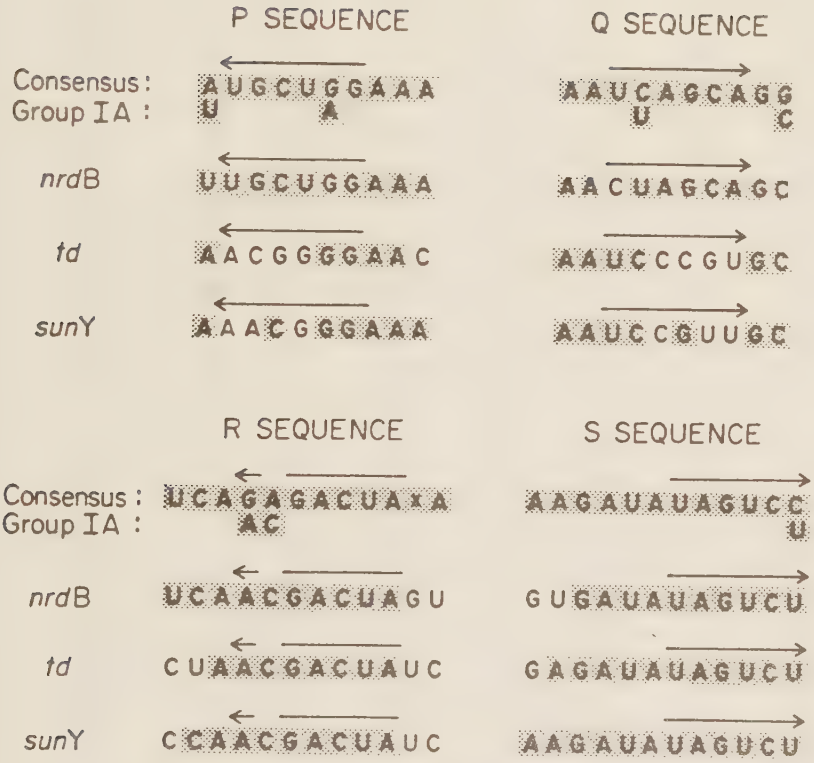


Abbildung 36.

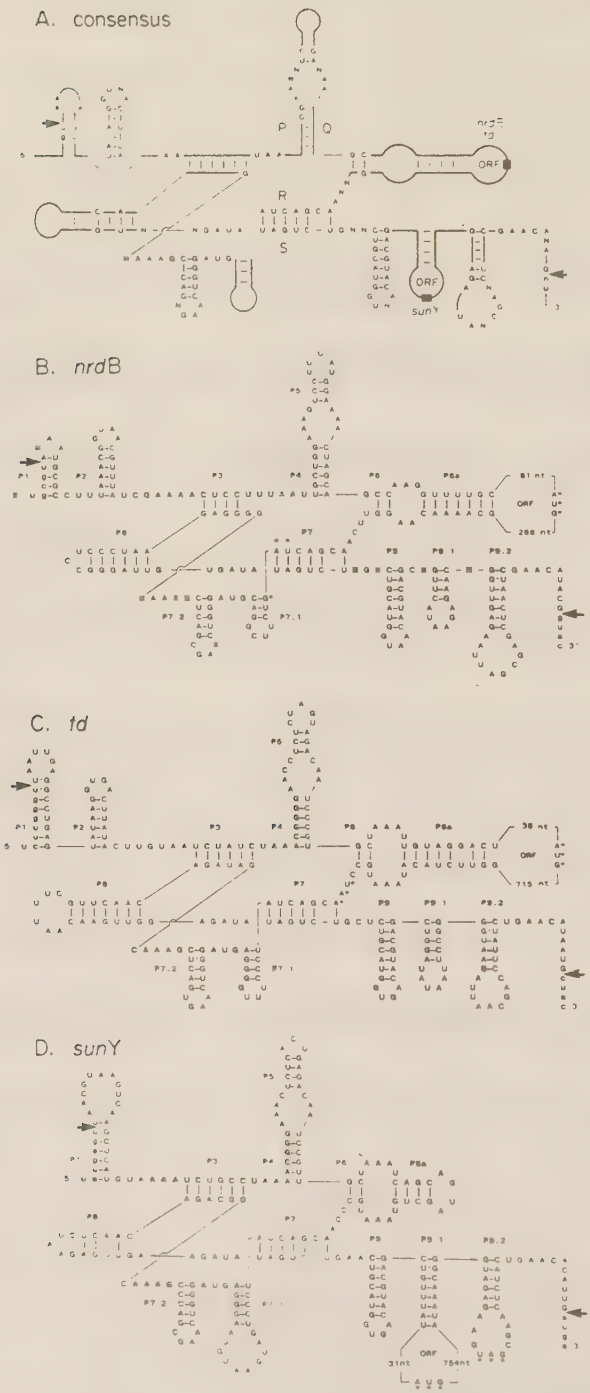


Abbildung 37.

Tabelle 11.

NAME	SEQUENCE	POSITION
ALUI	AGCT	44
SAU96I	GGGCC	170
HAEIII	GGCC	171
AHAIII	TTTAAA	198
SCAI	AGTACT	206
RSAI	GTAC	207
MAEI	CTAG	210
AHAIII	TTTAAA	232
MBOII	GAAGA	306
HPAII	CCGG	340
ESPI	GCTGAGC	373
DDEI	CTGAG	374
ALUI	AGCT	377
MAEI	CTAG	391
MNLI	CCTC	424
MSTII	CCTCAGG	424
DDEI	CTCAG	425
AHAIII	TTTAAA	484
MAEI	CTAG	519
BSTNI	CCTGG	542
SCRFI	CCTGG	542
NLAIII	CATG	566
ALUI	AGCT	578
BANI	GGTACC	645
KPNI	GGTACC	645
NLAIV	GGTACC	645
RSAI	GTAC	646
TAQI	TCGA	661
MBOII	GAAGA	690
AATII	GACGTC	699
AHAII	GACGTC	699
MAEIII	GTCAC	702
HPHI	TCACC	703
HINFI	GATTC	722
HAEIII	GGCC	732
NLAIV	GGCCCC	732
SAU96I	GGCCC	732
BSTNI	CCTGG	736
SCRFI	CCTGG	736
NLAIII	CATG	750
SSPI	AATATT	797
SFANI	GATGC	821
DDEI	CTGAG	830
MBOII	GAAGA	845
NLAIII	CATG	851
SFANI	GATGC	854
NSII	ATGCAT	855
TAQI	TCGA	879
TAQI	TCGA	906
MBOII	GAAGA	914
RSAI	GTAC	980
NSPHI	ACATGT	982
NLAIII	CATG	983
HINFI	GAATC	998
DDEI	CTCAG	1006
ALUI	AGCT	1062
TAQI	TCGA	1066
MAEIII	GTAAC	1079
DDEI	CTCAG	1083
RSAI	GTAC	1091
FNU4H	GCGGC	1126
MAEI	CTAG	1130
HPAII	CCGG	1156
MBOII	GAAGA	1159
HPAII	CCGG	1165
MBOII	GAAGA	1180
NLAIII	CATG	1237
HAEIII	GGCC	1299
HPAII	CCGG	1301
DDEI	CTTAG	1307
RSAI	GTAC	1317
AHAIII	TTTAAA	1325
MBOII	GAAGA	1355
SNABI	TACGTA	1438
HINFI	GAATC	1463
MBOII	GAAGA	1487
HPHI	GGTGA	1544
ALUI	AGCT	1567
ALUI	AGCT	1576
BBVI	GCTGC	1577
FNU4H	GCTGC	1577
ALUI	AGCT	1621
BBVI	GCTGC	1622
FNU4H	GCTGC	1622
FNU4H	GCCGC	1625
RSAI	GTAC	1633
HINFI	GATTC	1678
SFANI	GCATC	1699
MBOII	GAAGA	1792
SFANI	GATGC	1813
ALUI	AGCT	1869
HPHI	TCACC	1875
MNLI	CCTC	1878
TAQI	TCGA	1880
NLAIII	CATG	1885
NLAIII	CATG	1891
AVAII	GGACC	1903
SAU96I	GGACC	1903
BBVI	GCTGC	1951
FNU4H	GCTGC	1951
PSTI	CTGCAG	1952
AHAIII	TTTAAA	2046
SSPI	AATATT	2071
HPHI	TCACC	2076
MNLI	GAGG	2128
HAEIII	GGCC	2139
SFANI	GATGC	2145
MNLI	GAGG	2156
HINFI	GAATC	2173
ALUI	AGCT	2199
BBVI	GCAGC	2214
FNU4H	GCAGC	2214
ALUI	AGCT	2216
DDEI	CTTAG	2218
AHAIII	TTTAAA	2229
AHAIII	TTTAAA	2288
BSMI	GAATGC	2297
HINFI	GATTC	2310
DRAII	GGGACCT	2335
NLAIV	GGGACC	2335
AVAII	GGACC	2336
SAU96I	GGACC	2336
BSTXI	CCAAAAATATGG	2371
AVAII	GGTCC	2381
NLAIV	GGTCCC	2381
SAU96I	GGTCC	2381
HHAI	GCGC	2403
NLAIII	CATG	2406
NLAIII	CATG	2430
MBOII	TCTTC	2453
AHAIII	TTTAAA	2483
BSMI	GAATGC	2491
HINDIII	AAGCTT	2511
ALUI	AGCT	2512
MSTI	TGCGCA	2522
HHAI	GCGC	2523
MSTI	TGCGCA	2582
HHAI	GCGC	2583
HPHI	TCACC	2597
SCAI	AGTACT	2636
RSAI	GTAC	2637
MNLI	GAGG	2679
FOKI	GGATG	2681
SFANI	GATGC	2682
SFANI	GATGC	2706
ACCI	GTATAC	2745
SFANI	GATGC	2751
NSII	ATGCAT	2752
BSMI	GCATTC	2754
HINDIII	AAGCTT	2761
ALUI	AGCT	2762
HINCII	GTTAAC	2771
HPAI	GTTAAC	2771
ALUI	AGCT	2785
NLAIII	CATG	2789
DDEI	CTGAG	2822
ALUI	AGCT	2851
DDEI	CTAAG	2926
MBOII	GAAGA	2935
NLAIII	CATG	2954
HPHI	GGTGA	2957
MAEIII	GTGAC	2958
NSPHI	ACATGC	2961
NLAIII	CATG	2962
BBVI	GCAGC	2965
FNU4H	GCAGC	2965
NSPBII	CAGCTG	2966
PVUII	CAGCTG	2966
ALUI	AGCT	2967
MNLI	CCTC	2977
MNLI	GAGG	2997
HPHI	GGTGA	2999
HINDIII	AAGCTT	3028
ALUI	AGCT	3029
AHAIII	TTTAAA	3101
DDEI	CTAAG	3162
HINDIII	AAGCTT	3164
ALUI	AGCT	3165
HINDIII	AAGCTT	3179
ALUI	AGCT	3180
ALUI	AGCT	3196
DDEI	CTAAG	3198
MBOII	TCTTC	3211
HINFI	GATTC	3224
HINFI	GAATC	3233
HAEIII	GGCC	3238
SAU96I	GGCCC	3238
MBOII	GAAGA	3251
SSPI	AATATT	3270
HPAII	CCGG	3332
HINFI	GATTC	3335
MBOII	GAAGA	3341
DDEI	CTTAG	3359
AHAIII	TTTAAA	3391
ALUI	AGCT	3457
NLAIII	CATG	3461
SSPI	AATATT	3473
ALUI	AGCT	3511
AHAIII	TTTAAA	3514
HPAII	CCGG	3522
AHAIII	TTTAAA	3542
NSPBII	CAGCTG	3549
PVUII	CAGCTG	3549
ALUI	AGCT	3550
BBVI	GCTGC	3551
FNU4H	GCTGC	3551
MBOII	GAAGA	3575
FOKI	CATCC	3593
FOKI	CATCC	3626
FOKI	GGATG	3679
SFANI	GATGC	3680
BBVI	GCTGC	3683
FNU4H	GCTGC	3683
AHAIII	TTTAAA	3694
HINFI	GACTC	3840
MAEIII	GTAAC	3850
RSAI	GTAC	3866
HINFI	GATTC	3894
DDEI	CTAAG	3962
HINDIII	AAGCTT	3964
ALUI	AGCT	3965
MBOII	GAAGA	3982
SAU3A	GATC	4009
ALUI	AGCT	4031
XBAI	TCTAGA	4038
MAEI	CTAG	4039
MSTI	TGCGCA	4045
HHAI	GCGC	4046
FOKI	GGATG	4058
ALUI	AGCT	4080
DDEI	CTAAG	4082
ALUI	AGCT	4095
HINFI	GATTC	4162
TAQI	TCGA	4165
HGAI	GACGC	4190
MAEIII	GTAAC	4210
HINFI	GAATC	4218
MBOII	TCTTC	4221
MNLI	CCTC	4225
SFANI	GCATC	4229

FNUD2	CGCG	4233
HINFI	GATTC	4240
ALUI	AGCT	4283
HINDIII	AAGCTT	4316
ALUI	AGCT	4317
ALUI	AGCT	4356
BSMI	GCATTC	4390
AHAIII	TTTAAA	4435
SSPI	AATATT	4489
BSMI	GAATGC	4536
HPAII	CCGG	4568
TAQI	TCGA	4591
HINFI	GACTC	4593
TAQI	TCGA	4631
MAEIII	GTTAC	4663
RSAI	GTAC	4680
NLAIII	CATG	4696
MBOII	GAAGA	4708
SSPI	AATATT	4714
HAEII	AGCGCC	4746
HHAI	GCGC	4747
MAEIII	GTCAC	4775
MNLI	GAGG	4828
HAEIII	GGCC	4830
BBVI	GCTGC	4846
FNU4H	GCTGC	4846
NSPHI	ACATGT	4900
NLAIII	CATG	4901
RSAI	GTAC	4910
NLAIII	CATG	4915
HINFI	GACTC	4923
SAU3A	GATC	4969
MBOII	GAAGA	4994
NSPHI	ACATGT	4999
NLAIII	CATG	5000
FNUD2	CGCG	5029
NLAIII	CATG	5061
NSII	ATGCAT	5062
HINFI	GAATC	5095
MAEI	CTAG	5109
DDEI	CTAAG	5134
HPAII	CCGG	5155
NCII	CCGGG	5155
SCRFI	CCGGG	5155
DDEI	CTTAG	5216
HAEIII	GGCC	5237
SAU96I	GGCCC	5237
HPHI	TCACC	5252
SFANI	GATGC	5292
NSII	ATGCAT	5293
MAEIII	GTTAC	5331
MBOII	TCTTC	5337
MLUI	ACGCGT	5347
FNUD2	CGCG	5348
ALUI	AGCT	5360
DDEI	CTAAG	5362
HHAI	GCGC	5386
DDEI	CTAAG	5407
ALUI	AGCT	5413
ALUI	AGCT	5426
MAEI	CTAG	5428
MAEIII	GTAAC	5439
MBOII	TCTTC	5444
ECORV	GATATC	5468
MBOII	GAAGA	5488
ACCI	GTCTAC	5498
XBAI	TCTAGA	5532
MAEI	CTAG	5533
HPHI	TCACC	5571
ECORI	GAATTC	5662
NLAIII	CATG	5726
HINCII	GTTAAC	5737
HPAI	GTTAAC	5737
HAEIII	GGCC	5743
SAU96I	GGCCC	5743
BSTNI	CCTGG	5746
SCRFI	CCTGG	5746
FOKI	GGATG	5749
MBOII	TCTTC	5765
MAEIII	GTCAC	5770
FOKI	GGATG	5776
SCAI	AGTACT	5812
RSAI	GTAC	5813
BSMI	GAATGC	5820
MAEIII	GTAAC	5828
HGAI	GCGTC	5869
STUI	AGGCCT	5898
HAEIII	GGCC	5899
HPAII	CCGG	5912
BSTEII	GGTGACC	5917

HPHI	GGTGA	5917
MAEIII	GTGAC	5918
MBOII	GAAGA	5944
HPAII	CCGG	5986
ECORV	GATATC	5995
MBOII	GAAGA	6025
MAEI	CTAG	6040
MBOII	TCTTC	6093
ALUI	AGCT	6107
MNLI	GAGG	6116
SAU3A	GATC	6119
BBVI	GCAGC	6130
FNU4H	GCAGC	6130
NLAIII	CATG	6167
AVAII	GGTCC	6221
SAU96I	GGTCC	6221
ALUI	AGCT	6316
MAEIII	GTAAC	6335
ALUI	AGCT	6364
MBOII	TCTTC	6434
MAEI	CTAG	6438
ECORI	GAATTC	6443
HGAI	GCGTC	6454
ALUI	AGCT	6477
BSMI	GAATGC	6481
BBVI	GCTGC	6485
FNU4H	GCTGC	6485
DDEI	CTTAG	6491
ECORI	GAATTC	6536
BSTNI	CCTGG	6562
SCRFI	CCTGG	6562
FOKI	GGATG	6565
MAEIII	GTTAC	6577
BANII	GAGCTC	6595
BSP1286	GAGCTC	6595
HGIAI	GAGCTC	6595
SACI	GAGCTC	6595
ALUI	AGCT	6596
SAU3A	GATC	6655
NSPBII	CCGCTG	6661
FNUD2	CGCG	6685
HHAI	GCGC	6686
HINFI	GATTC	6692
HINFI	GATTC	6736
XBAI	TCTAGA	6739
MAEI	CTAG	6740
ECORV	GATATC	6799
MBOII	GAAGA	6808
APAI	GGGCCC	6820
BANII	GGGCCC	6820
BSP1286	GGGCCC	6820
DRAII	GGGCCCT	6820
NLAIV	GGGCCC	6820
SAU96I	GGGCC	6820
HAEIII	GGCC	6821
SAU96I	GGCCC	6821
SFII	GGCCCTCTGGGC	6821
BGLI	GCCCTCTGGGC	6822
MNLI	CCTC	6824
SAU96I	GGGCC	6829
HAEIII	GGCC	6830
DDEI	CTAAG	6864
ALUI	AGCT	6867
SSPI	AATATT	6897
NDEI	CATATG	6948
MNLI	GAGG	6964
MAEIII	GTTAC	7041
MAEIII	GTTAC	7085
SAU3A	GATC	7107
DDEI	CTAAG	7132
NCOI	CCATGG	7148
NLAIII	CATG	7149
ALUI	AGCT	7215
DDEI	CTCAG	7263
HINFI	GATTC	7294
FNUD2	CGCG	7330
HPHI	GGTGA	7333
MAEIII	GTTAC	7433
NSPBII	CAGCTG	7437
PVUII	CAGCTG	7437
ALUI	AGCT	7438
DDEI	CTGAG	7440
MNLI	GAGG	7442
NSPHI	GCATGC	7464
NSPHI	GCATGC	7464
SPHI	GCATGC	7464
NLAIII	CATG	7465
AHAIII	TTTAAA	7492
MAEIII	GTTAC	7499
AHAIII	TTTAAA	7578
HINFI	GAATC	7585
HINFI	GAATC	7592

BBVI	GCAGC	7606
FNU4H	GCAGC	7606
MBOII	GAAGA	7619
FNUD2	CGCG	7625
HHAI	GCGC	7626
NLAIII	CATG	7645
MNLI	GAGG	7664
FOKI	GGATG	7666
BBVI	GCAGC	7706
FNU4H	GCAGC	7706
NLAIII	CATG	7710
ALUI	AGCT	7720
RSAI	GTAC	7747
BANI	GGTGCC	7763
NLAIV	GGTGCC	7763
MSTII	CCTAAGG	7780
DDEI	CTAAG	7781
HAEIII	GGCC	7792
MNLI	GAGG	7825
NDEI	CATATG	7873
MAEI	CTAG	7896
MAEIII	GTAAC	7914
MNLI	CCTC	7920
ALUI	AGCT	7948
BCLI	TGATCA	7973
SAU3A	GATC	7974
MBOII	GAAGA	7979
HGAI	GACGC	7987
HINDIII	AAGCTT	8044
ALUI	AGCT	8045
HINFI	GAATC	8058
HINDIII	AAGCTT	8065
ALUI	AGCT	8066
NLAIII	CATG	8079
NLAIII	CATG	8108
DDEI	CTAAG	8152
ALUI	AGCT	8173
HINFI	GATTC	8253
FOKI	CATCC	8257
HINCII	GTTGAC	8321
HINFI	GATTC	8333
MAEI	CTAG	8339
SFANI	GATGC	8342
BBVI	GCAGC	8345
FNU4H	GCAGC	8345
ALUI	AGCT	8347
FNU4H	GCGGC	8357
HINFI	GATTC	8386
FOKI	CATCC	8390
MNLI	CCTC	8393
MBOII	TCTTC	8413
SSPI	AATATT	8419
BGLII	AGATCT	8435
XHOII	AGATCT	8435
SAU3A	GATC	8436
ALUI	AGCT	8449
DDEI	CTAAG	8451
MNLI	GAGG	8455
HPHI	GGTGA	8457
DDEI	CTAAG	8474
SNABI	TACGTA	8491
ALUI	AGCT	8507
ALUI	AGCT	8517
HINDIII	AAGCTT	8693
ALUI	AGCT	8694
NSPHI	ACATGC	8745
NLAIII	CATG	8746
HINDIII	AAGCTT	8762
ALUI	AGCT	8763
NSII	ATGCAT	8774
BSMI	GCATTC	8776
MAEIII	GTTAC	8828
MAEIII	GTGAC	8858
HGAI	GACGC	8919
MBOII	GAAGA	8929
HINFI	GAATC	8932
HPAII	CCGG	8950
NCII	CCGGG	8950
SCRFI	CCGGG	8950
HINFI	GATTC	8980
MBOII	TCTTC	8983
MAEI	CTAG	8987
MNLI	GAGG	9007
MNLI	CCTC	9022
HPAII	CCGG	9027
HINFI	GATTC	9030
MBOII	GAAGA	9038
ALUI	AGCT	9058
DDEI	CTCAG	9085

2.

1.		T			C			A			G			3.
		E.c.	T4		E.c.	T4		E.c.	T4		E.c.	T4		
T	Phe	26	69,5	Ser	36	36,3	Tyr	26	79,0	Cys	36	66,7	T	
	Phe	74	30,5	Ser	34	5,7	Tyr	74	21,0	Cys	64	33,3	C	
	Leu	3	32,9	Ser	2	22,3	-	-	-	-	-	-	A	
	Leu	4	14,5	Ser	5	5,1	-	-	-	Trp	100	100	G	
C	Leu	6	30,0	Pro	10	45,9	His	20	74,7	Arg	67	38,3	T	
	Leu	8	5,8	Pro	1	7,1	His	80	25,3	Arg	30	12,5	C	
	Leu	1	10,6	Pro	12	30,6	Gln	18	65,0	Arg	1	18,3	A	
	Leu	79	7,2	Pro	77	16,5	Gln	82	35,0	Arg	0,3	2,5	G	
A	Ile	20	70,0	Thr	36	54,3	Asn	8	74,9	Ser	4	20,4	T	
	Ile	79	11,9	Thr	48	13,6	Asn	92	25,1	Ser	19	10,2	C	
	Ile	1	18,2	Thr	5	21,4	Lys	71	74,7	Arg	1	22,5	A	
	Met	100	100	Thr	10	10,7	Lys	29	25,3	Arg	0,3	5,8	G	
G	Val	44	50,0	Ala	37	52,0	Asp	36	78,0	Gly	54	34,0	T	
	Val	9	10,5	Ala	10	9,9	Asp	64	22,0	Gly	42	16,3	C	
	Val	27	28,4	Ala	25	27,0	Glu	76	82,7	Gly	1	40,5	A	
	Val	19	11,1	Ala	28	11,2	Glu	24	17,3	Gly	3	9,2	G	

Tabelle 12.

2

1		T	E.c.	nrdC	pinA	49	sunY	55.12	55		C	E.c.	nrdC	pinA	49	sunY	55.12	55		3
		Phe	26	100	100	83,3	64,3	92,3	62,5		Ser	36	0	50,0	42,9	38,2	33,3	38,5	T	
		Phe	74	0	0	16,7	35,7	7,7	37.5		Ser	34	0	0	14,3	5,9	4,2	0	C	
	T	Leu	3	14,3	30,0	52,4	31,1	38,9	60,0		Ser	2	0	33,3	28,6	17,6	20,8	23,1	A	
		Leu	4	28,6	20,0	9,5	11,1	11.1	10,0		Ser	5	0	0	0	5,9	8,3	0	G	
		Leu	6	28,6	40,0	23,8	42,2	44,4	30,0		Pro	10	50,0	50,0	66,7	62,5	14,3	25,0	T	
		Leu	8	0	0	4,8	4,4	5,6	0		Pro	1	0	0	0	0	14,3	0	C	
	C	Leu	1	14,3	0	9,5	4,4	0	0		Pro	12	0	25,0	33,3	16,7	57,1	50,0	A	
		Leu	79	14,3	10,0	0	6,7	0	0		Pro	77	50,0	25,0	0	20,8	14,3	25,0	G	
		Ile	20	66,7	86,7	66,7	84,1	42,9	66,7		Thr	36	75,0	66,7	40,0	55,0	33,3	40,0	T	
		Ile	79	33,3	6,7	0	9,1	9,5	16,7		Thr	48	0	0	40,0	12,5	16,7	20,0	C	
	A	Ile	1	0	6,7	33,3	6,8	47,7	16,7		Thr	5	25,0	22,2	20,0	25,0	33,3	20,0	A	
		Met	100	100	100	100	100	100	100		Thr	10	0	11,1	0	7,5	16,7	20,0	G	
		Val	44	20,0	40,0	75,0	54,1	61,5	45,5		Ala	37	66,6	40,0	42,9	48,7	58,3	38,5	T	
		Val	9	0	20,0	0	5,4	7,7	18,2		Ala	10	0	20,0	14,3	5,4	25,0	7,7	C	
	G	Val	27	40,0	40,0	0	27,0	23,1	36,4		Ala	25	33,3	40,0	42,9	24,3	16.7	38,5	A	
		Val	19	40,0	0	25,0	13,5	7,7	0		Ala	28	0	0	0	21,7	0	15,4	G	

2

1		A	E.c.	nrdC	pinA	49	sunY	55.12	55		G	E.c.	nrdC	pinA	49	sunY	55.12	55		3
		Tyr	26	75,0	37,5	66,7	78,3	86,7	66,7		Cys	36	50,0	100	25,0	76,9	50,0	100	T	
		Tyr	74	25,0	62,5	33,3	21,7	13,3	33,3		Cys	64	50,0	0	75,0	23,1	50,0	0	C	
	T	-	-	-	-	-	-	-	-		-	-	-	-	-	-	-	-	A	
		-	-	-	-	-	-	-	-		Trp	100	0	100	100	100	100	100	G	
		His	20	100	100	60,0	62,5	87,5	50,0		Arg	67	66,7	42,8	50,0	30,4	12,5	83,3	T	
		His	80	0	0	40,0	37,5	12,5	50,0		Arg	30	0	0	0	21,7	0	0	C	
	C	Gln	18	50,0	72,8	80,0	52,9	66,7	71,4		Arg	1	0	42,8	16,7	13,0	12,5	0	A	
		Gln	82	50,0	22,2	20,0	47,1	33,3	28,6		Arg	0,3	33,3	0	0	4,3	12,5	0	G	
		Asn	8	33,3	80,0	83,3	68,4	89,5	69,2		Ser	4	50,0	16,7	14,3	20,6	29,2	23,1	T	
		Asn	92	66,7	20,0	16,7	31,6	10,5	30,8		Ser	19	50,0	0	0	11,8	4,2	7,7	C	
	A	Lys	71	88,9	55,6	63,2	87,0	82,6	73,7		Arg	1	0	14,3	33,3	21,7	62,5	16,7	A	
		Lys	29	11,1	44,4	36,8	13,0	17,4	26,3		Arg	0,3	0	0	0	8,7	12.5	0	G	
		Asp	36	85,7	80,0	62,5	78,8	85,7	68,8		Gly	54	57,1	25,0	25,0	43,2	71,4	16,7	T	
		Asp	64	14,3	20,0	37,5	21,2	14,3	31,3		Gly	42	14,3	25,0	16,7	6,8	0	16,7	C	
	G	Glu	76	60,0	84,2	84,6	82,1	80,0	85,7		Gly	1	28,6	37,5	58,3	45,5	28,6	33,3	A	
		Glu	24	40,0	15,8	15,4	17,9	20,0	14,3		Gly	3	0	12,5	0	4,5	0	33,3	G	

Tabelle 13.

Tab.14: Verzeichnis der in dieser Arbeit konstruierten rekombinanten Vektoren.

Klonbezeichnung	Besonderheiten
M13 JT36/2	: 0.6 kb-HaeIII-Fragment mit Gen nrdC
M13 JT37/21	: 1,6 kb-HpaII-PvuII-Fragment mit Gen 49 (+Sk-1) und
M13 JT37/25	: orf 55.13 in beiden Orientierungen
M13 JT41/3	: Gen 49 (-Sk-1), Rb-ATG: 7bp, ca. 700 bp großes Ins.
M13 JT41/8	: Gen 49 (-Sk-1), Rb-ATG: 5bp, ca. 1000 bp großes Ins.
M13 JT41/17	: Gen 49 (-Sk-1), Rb-ATG: 7bp, ca. 1000 bp großes Ins.
M13 JT41/19	: Gen 49 (-SK-1), Rb-ATG: 12bp, ca. 300 bp großes Ins.

Anm.: Die M13 Klone werden in dieser Arbeit häufig auch ohne den Zusatz "M13" genannt.

pJT7-6/nrdC	: Insert aus M13 JT36/2, Synthese von T4-Thioredoxin
pJT7-5/49	: Inserts aus M13 JT37/25 und M13 JT37/21, Überpro-
pJT7-6/49	: duktion von Genprodukt 55.13
pJT7-5/49A	: enthält Gen 49 (+Sk-1) und einen Teil von orf 55.13, Synthese eines etwa 12 kd großen Proteins
pJT7-6/orf55.13	: trägt orf 55.13 vollständig, Synthese von gp 55.13
pJT7-4/41-3-15M	: Insert von M13 JT41/3, kodiert wahrscheinlich aufgrund einer Mutation für ein inaktives gp 49
pJT7-4/41-3	: Insert von M13 JT41/3
pJT7-4/41-8	: Insert von M13 JT41/8
pJT7-4/41-17	: Insert von M13 JT41/17
pJT7-4/41-19	: Insert von M13 JT41/19
pJT7-6TpAT8	: pT7-6 mit einer poly(dA)·poly(dT) Sequenz zur Transkriptionstermination ausgehend von plasmidkodierten Promotorstrukturen
pJT7-6TpAT8/41-8-1	: pJT7-6TpAT8 mit dem Insert aus M13 JT41/8, synthetisiert aktive Endonuklease VII
pJT7-6TpAT8/41-8-4	: wie pJT7-6/pAT8/41-8-1

VERÖFFENTLICHUNGEN

J. Tomaschewski (1984):
Die Nukleotidsequenz des Genes für die ß-Glukosyltransferase des Bakteriophagen T4 und anliegender Regionen.
Diplomarbeit an der Abteilung Biologie der Ruhr-Universität Bochum

J. Tomaschewski, H. Gram, J.W. Crabb, W. Rüger (1985):
T4-induced alpha- and beta-glucosyltransferase: cloning of the genes and a comparison of their products based on sequencing data.
Nucl. Acids Res. 13: 7551-7568

H. Gram, J. Tomaschewski, W. Rüger (1985):
Cloning and Sequencing of T4 beta- and T2, T4 and T6 alpha-glucosyltransferases.
5th International T4 Evergreen Meeting, Olympia (USA), Abstract

J. Tomaschewski und W. Rüger (1987):
Nucleotide sequence and primary structures of gene products coded for by the T4 genome between map positions 48,266 kb and 39,166 kb.
Nucl. Acids Res. 15: 3632-3633

N. Lamm, J. Tomaschewski und W. Rüger (1987):
Nucleotide Sequence of the Deoxycytidylate Hydroxymethylase Gene of Bacteriophage T4 (g42) and the Homology of its Gene Product with Thymidylate Synthase of E. coli.
Nucl. Acids Res. 15: 3920

E. Kutter, B. Guttman, W. Rüger, J. Tomaschewski, G. Mosig (1987):
The Bacteriophage T4 Genomic Map.
In: Genetic maps 1987. A compilation of linkage and restriction maps of genetically studied organisms. Volume 4.
S.J. O´Brien (ed.), Laboratory of viral carcinogenesis NIH

J. Tomaschewski und W. Rüger (1987):
Difficulties on the way to over-expressing Endonuclease VII.
6th International T4 Evergreen Meeting, Olympia (USA), Abstract

K. Skorupski, J. Tomaschewski, W. Rüger und L.D. Simon (1987):
A T4 Gene that functions to inhibit E. coli Lon Protease.
6th International T4 Evergreen Meeting, Olympia (USA), Abstract

D.A. Shub, J.M. Gott, M.-Q. Xu, B.F. Lang, F. Michel, J. Tomaschewski, J. Pedersen-Lane, M. Belfort (1988):
Structure Conservation between three homologous Introns of Phage T4 and the group I Introns of Eukaryotes.
Proc. Natl. Acad. Sci. USA 85: 1151-1155

K. Skorupski, J. Tomaschewski, W. Rüger, L.D. Simon (1988):
A bacteriophage T4 gene which functions to inhibit Escherichia coli lon protease.
J. Bacteriol., zur Publikation eingereicht

L E B E N S L A U F

9.7.1958	: geboren in Dortmund
April 1966 - Juni 1969	: Besuch der Volksschule in Dortmund-Marten
Aug. 1969 - Juli 1970	: Besuch der Hauptschule Marten
Aug. 1970 - Juni 1977	: Besuch des Aufbaugymnasiums in Dortmund-Brünninghausen
Juni 1977	: Ablegung der Reifeprüfung
Okt. 1978 - Juni 1984	: Studium der Biologie an der Ruhr-Universität Bochum
Juni 1984	: Diplom im Fach Biologie an der Ruhr-Universität Bochum. Hauptfach: Mikrobiologie; Nebenfächer: Biochemie und Biophysik Organische Chemie
Seit Juni 1984	: Anfertigung der Dissertation am Lehrstuhl für Biologie der Mikroorganismen an der Ruhr-Universität Bochum